

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 March 2001 (29.03.2001)

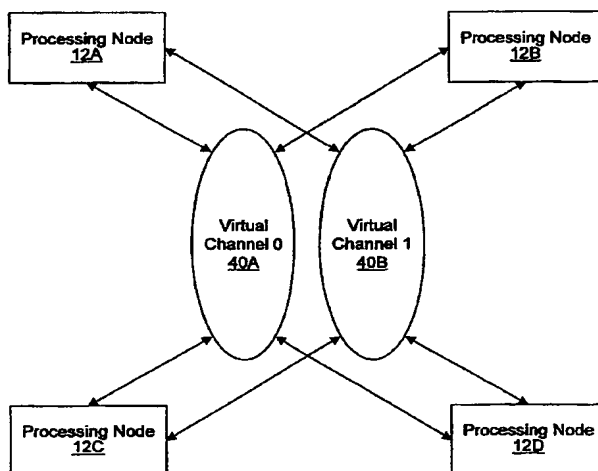
PCT

(10) International Publication Number
WO 01/22247 A1

- (51) International Patent Classification⁷: **G06F 15/173** (74) Agent: **RILEY, Louis, A.**; Advanced Micro Devices, Inc.,
5204 East Ben White Boulevard, M/S 562, Austin, TX
78741 (US).
- (21) International Application Number: **PCT/US00/12574**
- (22) International Filing Date: **9 May 2000 (09.05.2000)** (81) Designated States (*national*): **JP, KR.**
- (25) Filing Language: **English** (84) Designated States (*regional*): **European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).**
- (26) Publication Language: **English**
- (30) Priority Data:
09/399,281 17 September 1999 (17.09.1999) **US** Published:
— *With international search report.*
- (71) Applicant: **ADVANCED MICRO DEVICES, INC.** [US/US]; One AMD Place, Mail Stop 68, Sunnyvale, CA 94088-3453 (US).
For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

- (72) Inventors: **KELLER, James, B.**; 210 Iris Way, Palo Alto, CA 94303 (US). **MEYER, Derrick, R.**; 2728 Creek's Edge Parkway, Austin, TX 78733 (US).

(54) Title: **VIRTUAL CHANNELS AND CORRESPONDING BUFFER ALLOCATIONS FOR DEADLOCK-FREE COMPUTER SYSTEM OPERATION**



(57) Abstract: A computer system (10) employs virtual channels (40A-40B) and allocates different resources to the virtual channels. Packets which do not have logical/protocol-related conflicts are grouped into a virtual channel. Accordingly, logical conflicts occur between packets in separate virtual channels. The packets within a virtual channel may share resources (and hence experience resource conflicts), but the packets within different virtual channels may not share resources. Additionally, each virtual channel may be assigned control packet buffers (60A-60C, 220) and data packet buffers (60D-60E, 220). If a control packet which does not specify a data packet is received, no data packet buffer space is allocated. If a control packet which does specify a data packet is received, both control packet buffer space and data packet buffer space is allocated.

WO 01/22247 A1

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2003-510686

(P2003-510686A)

(43) 公表日 平成15年3月18日 (2003.3.18)

(51) Int.Cl. ⁷	識別記号	F I	テマコード* (参考)
G 0 6 F 15/177	6 7 6	G 0 6 F 15/177	6 7 6 A 5 B 0 0 5
	6 7 8		6 7 8 G 5 B 0 4 5
12/08	5 3 1	12/08	5 3 1 B 5 B 0 7 7
13/38	3 1 0	13/38	3 1 0 C 5 K 0 3 3
H 0 4 L 12/28	1 0 0	H 0 4 L 12/28	1 0 0 C
		審査請求 未請求	予備審査請求 有 (全 92 頁)

(21) 出願番号 特願2001-525542(P2001-525542)
(86) (22) 出願日 平成12年5月9日(2000.5.9)
(85) 翻訳文提出日 平成14年3月13日(2002.3.13)
(86) 国際出願番号 PCT/US00/12574
(87) 国際公開番号 WO01/022247
(87) 国際公開日 平成13年3月29日(2001.3.29)
(31) 優先権主張番号 09/399, 281
(32) 優先日 平成11年9月17日(1999.9.17)
(33) 優先権主張国 米国 (US)
(81) 指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), JP, KR

(71) 出願人 アドバンスト・マイクロ・ディバイズ・インコーポレイテッド
ADVANCED MICRO DEVICES INCORPORATED
アメリカ合衆国、94088-3453 カリフォルニア州、サニベイ、ピィ・オー・ボックス・3453、ワン・エイ・エム・ディ・プレイス、メール・ストップ・68 (番地なし)

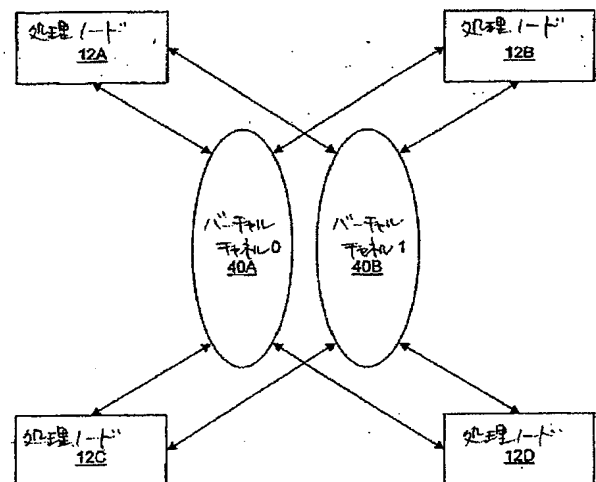
(74) 代理人 弁理士 深見 久郎 (外5名)

最終頁に続く

(54) 【発明の名称】 デッドロックのないコンピュータシステム動作のためのバーチャルチャネルおよび対応するバッファ割当て

(57) 【要約】

コンピュータシステム (10) はバーチャルチャネル (40A~40B) を用い異なる資源をバーチャルチャネルに割当てる。ロジック的/プロトコルに関連するコンフリクトを有さないパケットがバーチャルチャネル内へグループ化される。したがってロジックコンフリクトは別々のバーチャルチャネルにおけるパケット間で生ずる。バーチャルチャネル内のパケットは資源を共有し (したがって資源コンフリクトを経験し) てもよいが、異なるバーチャルチャネル内のパケットは資源を共有してはならない。加えて、各バーチャルチャネルは制御パケットバッファ (60A~60C、220) およびデータパケットバッファ (60D~60E、220) を割当てられてもよい。データパケットを特定しない制御パケットが受信される場合、データパケットバッファ空間は全く割当てられない。データパケットを特定する制御パケットが受信される場合、制御パケットバッファ空間およびデータパケットバッファ空間の両方が割当てられる。



【特許請求の範囲】

【請求項1】 パケットをコンピュータシステム（10）において複数のノード（12A～12D、204A～204C）間で経路づけるための方法であって、

第1の制御パケットを前記複数のノード（12A～12D、204A～204C）のうち第1のノード（12A、204A）において受取るステップを含み、前記第1のノード（12A、204A）は複数の制御パケットバッファ（60A～60C、220）を含み、前記複数の制御パケットバッファ（60A～60C、220）の各々は複数のバーチャルチャネル（40A～40B）のうち異なるバーチャルチャネルに割当てられ、前記方法はさらに、

前記複数のバーチャルチャネル（40A～40B）のうち前記第1の制御パケットが属する第1のバーチャルチャネル（40A）を判断するステップと、

前記第1の制御パケットを前記複数の制御パケットバッファ（60A～60C、220）のうち第1の制御パケットバッファ（60A、220）において記憶し、前記第1の制御パケットバッファ（60A、220）は前記第1のバーチャルチャネル（40A）に割当てられる、方法。

【請求項2】 前記第1の制御パケットにより特定される第1のデータパケットを受取るステップと、

前記第1のデータパケットを前記第1のノード（12A、204A）内において複数のデータバッファ（60D～60E、220）のうち第1のデータバッファ（60D、220）において記憶し、前記複数のデータバッファ（60D～60E、220）の各々は、前記複数のバーチャルチャネル（40A～40B）のうち、対応するデータパケットを特定する少なくとも1つの制御パケットを含む異なるバーチャルチャネルに割当てられる、請求項1に記載の方法。

【請求項3】 第2の制御パケットを前記第1のノード（12A、204A）において受取るステップと、

前記複数のバーチャルチャネル（40A～40B）のうち前記第2の制御パケットが属する第2のバーチャルチャネル（40B）を判断するステップとを含み、前記第2のバーチャルチャネル（40B）は前記第1のバーチャルチャネル（

40A)とは異なり、前記方法はさらに、

前記第2の制御パケットを前記複数の制御パケットバッファ(60A~60C、220)のうち第2の制御パケットバッファ(60B、220)において記憶するステップを含み、前記第2の制御パケットバッファ(60B、220)は前記第2のバーチャルチャネル(40B)に割当てられる、請求項1に記載の方法。

【請求項4】 前記第1の制御パケットを前記複数のノード(12A~12D、204A~204C)のうち第2のノード(12B、204B)に送信し、前記第2のノード(12B、204B)は第2の複数の制御パケットバッファ(60A~60C、220)を含み、前記第2の複数の制御パケットバッファ(60A~60C、220)の各々は前記複数のバーチャルチャネル(40A~40B)のうち異なるバーチャルチャネルに割当てられ、前記送信は前記第2の複数の制御パケットバッファ(60A~60C、220)のうち前記第1の制御パケットを記憶する空間を含む第3の制御パケットバッファ(60A、220)に応答し、前記第3の制御パケットバッファ(60A、220)は前記第1のバーチャルチャネル(40A)に割当てられ、前記方法はさらに、

前記第2の複数の制御パケットバッファ(60A~60C、220)のうち前記第2の制御パケットを記憶する空間を含む第4の制御パケットバッファ(60B、220)に応答して前記第2の制御パケットを前記第2のノードに送信するステップを含み、前記第4の制御パケットバッファ(60B、220)は前記第2のバーチャルチャネル(40B)に割当てられる、請求項3に記載の方法。。

【請求項5】 第1の制御パケットを送信するよう構成される第1のノード(12A、204B)と、

前記第1の制御パケットを前記第1のノード(12A、204B)から受信するよう結合される第2のノード(12B、204A)とを含み、前記第2のノード(12B、204A)は複数の制御パケットバッファ(60A~60C、220)を含み、前記複数の制御パケットバッファ(60A~60C、220)の各々は複数のバーチャルチャネル(40A~40B)のうち異なるバーチャルチャネルに割当てられ、前記第2のノード(12B、204A)は、前記複数のバー

チャルチャネル（４０Ａ～４０Ｂ）のうち前記第１の制御パケットが属する第１のバーチャルチャネル（４０Ａ）に応答して前記複数の制御パケットバッファ（６０Ａ～６０Ｃ、２２０）のうち第１の制御パケットバッファ（６０Ａ、２２０）において前記第１の制御パケットを記憶するよう構成される、コンピュータシステム。

【請求項６】 前記第２のノード（１２Ｂ、２０４Ａ）は複数のデータバッファ（６０Ｄ～６０Ｅ、２２０）をさらに含み、前記複数のデータバッファ（６０Ｄ～６０Ｅ、２２０）の各々は、前記複数のバーチャルチャネル（４０Ａ～４０Ｂ）のうち、対応するデータパケットを特定する少なくとも１つの制御パケットを含む、異なるバーチャルチャネルに割当てられる、請求項５に記載のコンピュータシステム。

【請求項７】 前記第１のノード（１２Ａ、２０４Ｂ）は前記第１の制御パケットにより特定される第１のデータパケットを送信するよう構成され、前記第２のノード（１２Ｂ、２０４Ａ）は前記第１のデータパケットを前記複数のデータバッファ（６０Ｄ～６０Ｅ、２２０）のうち第１のデータバッファ（６０Ｄ、２２０）に記憶するよう構成され、前記第１のデータバッファ（６０Ｄ、２２０）は前記第１のバーチャルチャネル（４０Ａ）に割当てられる、請求項６に記載のコンピュータシステム。

【請求項８】 前記第１のノード（１２Ａ、２０４Ｂ）は前記複数のバーチャルチャネル（４０Ａ～４０Ｂ）のうち第２のバーチャルチャネル（４０Ｂ）に属する第２の制御パケットを送信するよう構成され、前記第２のバーチャルチャネル（４０Ｂ）は前記第１のバーチャルチャネル（４０Ａ）とは異なり、前記第２のノード（１２Ｂ、２０４Ａ）は前記第２の制御パケットを前記複数の制御パケットバッファ（６０Ａ～６０Ｃ、２２０）のうち第２の制御パケットバッファ（６０Ｂ、２２０）において記憶するよう構成される、請求項５に記載のコンピュータシステム。

【請求項９】 前記第２のノード（１２Ｂ、２０４Ａ）が前記第１の制御パケットのデスティネーションである場合、前記第２のノード（１２Ｂ、２０４Ａ）は前記第１の制御パケットを前記第１の制御パケットバッファ（６０Ａ、２２

0) から取除き前記第1の制御パケットに応答するよう構成され、前記第2のノード(12B、204A)は前記第1のデータパケットを前記第1のデータバッファ(60D、220)から取除き前記第1のデータパケットを処理するようさらに構成される、請求項7に記載のコンピュータシステム。

【請求項10】 パケットを前記第2のノード(12B、204A)から受信するよう結合される第3のノード(12C、204C)をさらに含み、前記第2のノード(12B、204A)が前記第1の制御パケットのデスティネーションでない場合、前記第2のノード(12B、204A)は前記第1の制御パケットを前記第1の制御パケットバッファ(60A、220)から取除き前記第1の制御パケットを前記第3のノード(12C、204C)へ送るよう構成され、前記第2のノード(12B、204A)は前記第1のデータパケットを前記第1のデータバッファ(60D、220)から取除き前記第1のデータパケットを前記第3のノード(12C、204C)に送るようさらに構成される、請求項7に記載のコンピュータシステム。

【発明の詳細な説明】**【0001】****【発明の背景】****1. 技術分野**

この発明はコンピュータシステムの分野に関し、より特定的には、コンピュータシステムにおけるノード間の相互接続に関する。

【0002】**2. 背景技術**

一般的には、パーソナルコンピュータ（PC）およびその他の種類のコンピュータシステムは、メモリにアクセスするために共用バスシステムを中心に設計されてきた。1つ以上のプロセッサおよび1つ以上の入力／出力（I／O）装置が、共用バスを介してメモリに結合される。I／O装置は、共用バスとI／O装置との間の情報の転送を管理するI／Oブリッジを介して共用バスに結合される場合もある一方、プロセッサは典型的には、直接共用バスに結合されるか、またはキャッシュ階層構造を介して共用バスに結合される。

【0003】

残念ながら、共用バスシステムはいくつかの欠点を有する。たとえば、共用バスには多数の装置が装着されることから、バスは典型的には比較的低い周波数で動作される。多数の装着物は、バス上で信号を駆動する装置に容量的な高負荷をもたらし、多数の装着点は、比較複雑な高周波数に対する伝送ラインモデルをもたらす。したがって、周波数は低く留められ、共用バスで利用できる帯域幅も同様に比較的低い。低帯域幅は共用バスに付加的な装置を装着するのに障壁になり得るが、これは利用できる帯域幅によって性能が制限されるおそれがあるためである。

【0004】

共用バスシステムの他の欠点は、より多くの装置に対するスケーラビリティの欠如である。上述のように、帯域幅が固定される（そして、もし付加的な装置の追加によってバスの動作可能周波数が減じられると、減少し得る）。バスに（直接的にまたは間接的に）装着された装置の帯域幅要件が、一旦バスの利用可能な

帯域幅を超えると、装置はバスへのアクセスを試みたときにしばしばストールし得る。全体的な性能が減じられるであろう。

【0005】

上述の問題のうちの1つまたはいくつかには、分散メモリシステムを用いて対処し得る。分散メモリシステムを用いるコンピュータシステムは、複数のノードを含む。2つ以上のノードがメモリに接続され、それらのノードは何らかの好適な相互接続を用いて相互接続される。たとえば、ノードの各々は専用ラインを用いて他のノードに互いに接続されることができ、これに代えて、ノードの各々は固定された数の他のノードに接続され、トランザクションは第1のノードから1つ以上の中間ノードを介して、第1のノードに直接接続されていない第2のノードに経路制御されてもよい。メモリアドレス空間は、各々のノードのメモリにわたって割当てられる。一般的に、「ノード」とは相互接続上のトランザクションに関わることでできる装置である。たとえば、パケットベースの相互接続においては、ノードは他のノードとパケットを送受信し得るよう構成される。特定のトランザクションを行なうために、1つ以上のパケットが採用され得る。特定のノードは、パケットに対するデスティネーションであり得るが、この場合情報はノードによって受取られて、ノード内で処理される。これに変えて、特定のノードは、もし特定のノードがパケットのデスティネーションノードでなければ、パケットをソースノードからデスティネーションノードへ中継するためにも用い得る。

【0006】

分散メモリシステムは、共用バスシステムの困難とは異なった設計上の困難を呈示する。たとえば、共用バスシステムはバスアービトレーションを介してトランザクションの開始を規制する。したがって、明確なアービトレーションアルゴリズムが、関与するバスにトランザクションを開始させる機会を与える。バス上のトランザクションの順序が、そのトランザクションが行なわれる順序を表し得る（たとえば、コヒーレンシ目的で）。一方、分散システムにおいては、ノードは並行にトランザクションを開始し、相互接続を用いてトランザクションを他のノードに送信し得る。これらのトランザクションは、それらの間でロジック的コ

ンフリクト（たとえば、同じアドレスへのトランザクションに対するコヒーレンスコンフリクト）を有し、資源コンフリクトが生じるおそれがあるが、これはトランザクションの開始を規制するための中央機構が提供されていないためである。したがって、情報がノード間を円滑に伝播しつづけ、かつ（トランザクション間のコンフリクトによりトランザクションが完了しない）デッドロック状態が回避されることを確実にすることが、より困難である。分散システムにおけるデッドロックを回避するための方法および装置が望まれる。さらに、実現を容易にするために、（ハードウェアの意味での）装置を最小化することが望まれる。

【0007】

【発明の開示】

上に概要を述べた問題は、ここで説明するようにバーチャルチャネルを用い、かつバーチャルチャネルに異なった資源を割当てることにより、大部分が解決される。ロジック／プロトコル関連のコンフリクトを有さないパケットは、バーチャルチャネルにグループ化される。したがって、ロジック的コンフリクトは別々のバーチャルチャネル内のパケットの間で生じる。1つのバーチャルチャネル内のパケットは資源を共用する（よって資源コンフリクトが生じる）が、異なったバーチャルチャネル内のパケットは資源を共用しないであろう。資源コンフリクトが生じるパケットはロジック的コンフリクトは生じなく、かつロジック的コンフリクトが生じるパケットは資源コンフリクトは生じないために、デッドロックのない動作が達成される。

【0008】

さらに、各バーチャルチャネルには、制御パケットバッファとデータパケットバッファとが割当てられ得る。制御パケットは、データパケットよりも実質的にサイズが小さく、データパケットよりも頻繁に生じ得る。別々のバッファを与えることにより、バッファ空間を効率的に用いることができる。データパケットを特定しない制御パケットが受取られると、データパケットバッファ空間は割当てられない。データパケットを特定する制御パケットが受取られると、制御パケットバッファ空間およびデータパケットバッファ空間の両方が割当てられる。制御パケットはデータパケットよりも小さくより頻繁に生じることが多いので、全体

的なバッファ記憶を実質的に増加させることなく、データパケットバッファ内よりも制御パケットバッファ内でより多くのバッファエントリを提供することができる。しかしながら、パケットのスループットは上昇し得る。

【0009】

概略的には、コンピュータシステムにおける複数のノード間のパケットの経路制御に対する方法が企図される。第1の制御パケットは、複数のノードのうちの第1のノードで受取られる。第1のノードは、複数の制御パケットバッファを含み、その各々が複数の異なったバーチャルチャネルに割当てられる。第1の制御パケットが属する、複数のバーチャルチャネルのうちの第1のバーチャルチャネルが決定される。第1の制御パケットは、複数の制御パケットバッファのうちの第1の制御パケットバッファに記憶され、第1の制御パケットバッファは第1のバーチャルチャネルに割当てられる。

【0010】

さらに、第1のノードおよび第2のノードを含むコンピュータシステムが企図される。第1のノードは、第1の制御パケットを送信するよう構成される。第1のノードからの第1の制御パケットを受けよう結合される第2のノードは複数の制御パケットバッファを含む。複数の制御パケットバッファの各々は、複数のバーチャルチャネルの異なった1つに割当てられる。第2のノードは、第1の制御パケットが属する複数のバーチャルチャネルの第1のバーチャルチャネルに応答する複数の制御パケットバッファの第1の制御パケットバッファにおける第1の制御パケットを記憶するよう構成される。

【0011】

この発明の他の目的と利点とは、以下の詳細な説明を読み、添付の図面を参照することにより、より明らかとなるであろう。

【0012】

この発明はさまざまな変形と代替形に対処するものであるが、その特定の実施例を図面において例示の目的で示し、以下に詳述する。しかしながら、図面とその詳細な説明とは開示される発明を特定の形に限定することを意図せず、反対に、すべての変形、等価物、および代替例は前掲の特許請求の範囲に規定されるこ

の発明の精神および範囲に入ること意図する。

【0013】

【発明の実施の形態】

システムの概要

図1は、マルチプロセッシングコンピュータシステム10の一実施例を示す。他の実施例が可能であり企図される。図1の実施例においては、コンピュータシステム10はいくつかの処理ノード12A、12B、12C、および12Dを含む。処理ノードの各々は、処理ノード12A-12Dにそれぞれ含まれるメモリコントローラ16A-16Dを介して、それぞれのメモリ14A-14Dに結合される。さらに、処理ノード12A-12Dは、処理ノード12A-12Dの間の通信のために用いるインターフェイスロジックを含む。たとえば、処理ノード12Aは、処理ノード12Bと通信するためのインターフェイスロジック18Aと、処理ノード12Cと通信するためのインターフェイスロジック18Bと、さらに別の処理ノード（図示せず）と通信するための第3のインターフェイスロジック18Cを含む。同様に、処理ノード12Bはインターフェイスロジック18D、18E、および18Fを含み、処理ノード12Cはインターフェイスロジック18G、18H、および18Iを含み、処理ノード12Dはインターフェイスロジック18J、18K、および18Lを含む。処理ノード12Dは、インターフェイスロジック18Lを介して結合されてI/Oブリッジ20と通信する。他の処理ノードは同様の様式で他のI/Oブリッジと通信し得る。I/Oブリッジ20はI/Oバス22に結合される。

【0014】

処理ノード12A-12Dは、処理ノード間通信のためのパケットベースのリンクを実現化する。この実施例においては、リンクは単方向ラインの組として実現化される（たとえば、ライン24Aはパケットを処理ノード12Aから処理ノード12Bへ送信するために用いられ、ライン24Bはパケットを処理ノード12Bから処理ノード12Aへ送信するために用いられる）。ライン24C-24Hの他の組は、図1に示すようにパケットを他の処理ノード間で通信するために用いられる。リンクは、処理ノード間の通信のためにキャッシュコヒーレント様

式で動作するか（「コヒーレントリンク」）、または処理ノードとI/Oブリッジとの間の通信のために非コヒーレント様式で動作し得る（「非コヒーレントリンク」）。さらに、非コヒーレントリンクは、I/Oバス22の代わりにI/O装置の間でデイジーチェーン構造として用いられ得る。コヒーレントリンクを介した2つ以上のノードの相互接続は、「コヒーレントファブリック」と呼ばれる。同様に、非コヒーレントリンクを介した2つ以上のノードの相互接続は、「非コヒーレントファブリック」と呼ばれる。一方の処理ノードから他方へ送信されるべきパケットは、1つ以上の中間ノードを通過し得ることに留意されたい。たとえば、処理ノード12Aから処理ノード12Dに送信されるパケットは、図1に示すように、処理ノード12Bまたは処理ノード12Cのいずれかを通過し得る。いかなる好適な経路制御アルゴリズムをも用い得る。コンピュータシステム10の他の実施例は、図1に示す実施例よりも多いかまたは少ない処理ノードを含み得る。

【0015】

処理ノード12A-12Dは、メモリコントローラとインターフェイスロジックに加えて、1つ以上のプロセッサを含み得る。概略的には、処理ノードは少なくとも1つのプロセッサを含み、選択により、メモリおよび所望の他のロジックとの間で通信するためのメモリコントローラを含む。

【0016】

メモリ14A-14Dは、何らかの好適なメモリ装置を含み得る。たとえば、メモリ14A-14Dは、1つ以上のRAMBUS DRAM（RDRAM）、シンクロナスDRAM（SDRAM）、スタティックRAMなどを含み得る。コンピュータシステム10のアドレス空間は、メモリ14A-14Dの間で分割される。処理ノード12A-12Dの各々はメモリマップを含むことができ、該メモリマップを用いて、どのアドレスがどのメモリ14A-14Dにマッピングされているか、よって、ある特定のアドレスに対するメモリ要求がどの処理ノード12A-12Dに経路制御されるべきかを判断する。一実施例においては、コンピュータシステム10内のアドレスに対するコヒーレンシ点は、アドレスに対応するバイトを記憶しているメモリに結合された、メモリコントローラ16A-1

6 Dである。言い換えると、メモリコントローラ16 A-16 Dは、対応するメモリ14 A-14 Dへのメモリアクセスの各々を、キャッシュコヒーレントな様式で起こることを確実にすることを担当している。メモリコントローラ16 A-16 Dは、メモリ14 A-14 Dにインターフェイスするための制御回路を含み得る。さらに、メモリコントローラ16 A-16 Dは、メモリ要求を待ち行列として管理するための、要求待ち行列を含み得る。

【0017】

一般的には、インターフェイスロジック18 A-18 Lは、リンクからパケットを受取り、かつリンクに送信されるべきパケットをバッファするための、バッファを含み得る。コンピュータシステム10は、パケットを送るために好適などのようなフロー制御をも用い得る。たとえば一実施例においては、ノードの各々は、インターフェイスロジックの各々が接続されたリンクの他端のレシーバ内に、いくつかの各種のバッファのカウントを記憶する。ノードは、受信ノードがパケットを記憶するための空いたバッファを有さない限り、パケットを送信しない。パケットを次に経路制御することにより受信バッファが空くと、受信インターフェイスロジックは送信インターフェイスロジックにメッセージを送り、バッファが空いたことを示す。そのような機構は、「クーポンに基づく」システムと呼ばれる。

【0018】

次に図2は、処理ノード12 Aおよび12 Bのブロック図を示し、それらの間のリンクの一実施例を詳細に例示する。他の例も可能であり、企図される。図2の実施例においては、ライン24 Aは、クロックライン24 A Aと、制御ライン24 A Bと、制御/アドレス/データバス24 A Cとを含む。同様に、ライン24 Bは、クロックライン24 B Aと、制御ライン24 B Bと、制御/アドレス/データバス24 B Cとを含む。

【0019】

クロックラインは、制御ラインおよび制御/アドレス/データバスに対するサンプルポイントを示すクロック信号を送信する。特定の一実施例においては、データ/制御ビットはクロック信号のエッジの各々（すなわち立上がりエッジおよ

び立下がりエッジ)で送信される。したがって、クロックサイクルごとに、ラインごとに2つのデータビットを送信し得る。ラインごとに1ビットを送信するために使用される時間は、ここでは「ビット時間」と呼ぶ。上述の実施例は、クロックサイクルごとに2つのビット時間を含む。パケットは2つ以上のビット時間で送信し得る。制御／アドレス／データバスの幅に依存して、多数のクロックラインを用い得る。たとえば32ビット制御／アドレス／データバスに対しては2つのクロックラインを用い得る(制御／アドレス／データバスの半分では一方のクロックラインが参照され、残りの半分の制御／アドレス／データバスと制御ラインとでは他方のクロックラインが参照される)。

【0020】

制御ラインは、制御／アドレス／データバスに送信されたデータが、ビット時間の制御パケットか、またはビット時間のデータパケットであるかを示す。制御ラインはアサートされてビット時間の制御パケットを示し、デアサートされてビット時間のデータパケットを示す。ある制御パケットは、後にデータパケットが続くことを示す。データパケットは、対応する制御パケットのすぐ後に続き得る。一実施例においては、他の制御パケットがデータパケットの送信に割込むおそれがある。そのような割込は、データパケットの送信の間に制御ラインをいくつかのビット時間アサートし、かつ制御ラインがアサートされている間にビット時間の制御パケットを送信することにより行なわれる可能性がある。データパケットに割込む制御パケットは、データパケットが後に続くことを示さないおそれがある。さらに、一実施例においては、制御ラインが制御パケットの送信の間にデアサートされてストールビット時間を示し得る。後の制御ラインの再アサートは、制御パケットが続いていることを示し得る。

【0021】

制御／アドレス／データバスは、データ／制御ビットを送信するための1組のラインを含む。一実施例においては、制御／アドレス／データバスは、8、16、または32のラインを含み得る。処理ノードまたはI/Oブリッジの各々は、設計選択にしたがってサポートされる数のラインのうちのいずれかを用い得る。他の実施例は、所望により他のサイズの制御／アドレス／データバスをサポート

し得る。

【0022】

一実施例によると、コマンド／アドレス／データバスラインおよびクロックラインは、反転データを担持し得る（すなわち、ロジック1はライン上の低電圧として表わされ、ロジック0が高電圧として表わされる）。これに代えて、これらのラインは非反転データを担持してもよい（ロジック1はライン上の高電圧として表わされ、ロジック0は低電圧として表わされる）。

【0023】

図3から図6は、コヒーレントリンクの一実施例において用いられる、例示的なパケットを示す。図3から図5は制御パケットを示し、図6はデータパケットを示す。他の実施例は、所望により異なったパケット定義を用い得る。パケットの各々は、「ビット時間」の見出しの下に列挙される一連のビット時間で示される。パケットのビット時間は、リストされたビット時間順序に従って送信される。図3から図6は、8ビット制御／アドレス／データバス実現化のためのパケットを示す。したがって、ビット時間の各々は、7から0まで番号が付与された8つのビットを含む。図中、いずれの値も付与されていないビットは、所与のパケットのために予約されているか、またはパケット特定情報を送信するために用いられるかのいずれかであり得る。点線で示されるフィールドは、あるタイプのすべてのパケットに含まれない可能性のある任意のフィールドを示す。

【0024】

一般的に、パケットは2つのノード（パケットを送信する起動ノードおよびパケットを受取るデスティネーションノード）間の通信である。起動ノードおよびデスティネーションノードは、パケットが関わるトランザクションのソースおよびターゲットノードとは異なるか、またはいずれかのノードがソースノードまたはターゲットノードのいずれかであり得る。制御パケットは、トランザクションに関する制御情報を担持するパケットである。ある制御パケットは、データパケットが後に続くことを特定する。データパケットは、トランザクションに対応し、かつ特定する制御パケットに対応するデータを担持する。

【0025】

図3は情報パケット（infoパケット）30を示す。情報パケット30は、8ビットリンク上の4つのビット時間を含む。この実施例においては、コマンド符号化はビット時間1の間に送信され、かつ6ビットを含む。図4、および図5に示す他方の制御パケットの各々は、ビット時間1の間に同じビット位置においてコマンド符号化を含む。メッセージがメモリアドレスを含まないときに、情報パケット30を用いてこのメッセージを処理ノード間で送信し得る。さらに、クーポンに基づくフロー制御方策を用いてバッファ空きカウントを送信するためにも、情報パケットを用い得る。

【0026】

図4はコマンドパケット（commandパケット）32を示す。コマンドパケット32は、8ビットリンク上の8つのビット時間を含む。コマンド符号化は、ビット時間1の間に送信される。ソースユニット番号も同様にビット時間1の間に送信され、ソースノード番号はビット時間2の間に送信される。ノード番号はコンピュータシステム10内の処理ノード12A-12Dのうちの1つを明確に識別し、かつパケットをコンピュータシステム10を介して経路制御するために用いられる。ユニット番号は、トランザクションを発したノード内のユニットを識別するか（ソースユニット番号）、またはトランザクションのデスティネーションであるノード内のユニットを識別する（デスティネーションユニット番号）。ユニットは、メモリコントローラ、キャッシュ、プロセッサなどを含み得る。任意で、コマンドパケット32はビット時間2におけるデスティネーションノード番号およびデスティネーションユニットのいずれかを含み得る（または他のいくつかのパケットでは、ターゲットノード番号およびターゲットユニット）。もしデスティネーション番号が含まれていれば、これはパケットをデスティネーションノードに経路制御するために用いられる。また、多くのコマンドパケットがビット時間3にソースタグを含み得るが、これはソースノードおよびソースユニットと併せて、パケットを関連する特定のトランザクションにリンクし得る。ビット時間5から8は、トランザクションによって影響されるメモリアドレスの最上位のビットを送信するために用いられる。コマンドパケット32は、トランザクションを開始し（たとえば、読出または書込トランザクション）、かつトランザク

ションによって影響されるメモリアドレスを担持するコマンドに対するトランザクションを実行するプロセスにおいてコマンドを送信するために用いられる。一般的に、コマンドパケットはデスティネーションノードによって実行されるべき動作を示す。

【0027】

パケット32内の定義されないいくつかのフィールドをコマンドパケットにおいて用いて、パケット特定情報を担持し得る。さらに、ビット時間4をいくつかのコマンド内で用いて、トランザクションによって影響されるメモリアドレスの最下位のビットを送信する。

【0028】

図5は、応答パケット（responseパケット）34を示す。応答パケット34は、コマンド符号化と、デスティネーションノード番号およびデスティネーションユニット番号とを含む。デスティネーションノード番号は、応答パケットに対するデスティネーションノードを識別する（これは、場合によってトランザクションのソースノードまたはターゲットノードであり得る）。さまざまな種類の応答パケットは付加的な情報を含み得る。たとえば、読出応答パケットは、後に続くデータパケットで提供される読出データの量を示し得る。プローブ応答は、プローブノードによって要求されたキャッシュブロックのコピーが保持されているか否かを示し得る（ビット時間4における任意の共用ビット“Sh”を用いる）。一般的に、応答パケット34は、トランザクションを行なう間に、トランザクションによって影響されるメモリアドレスの送信を必要としないコマンドに対して用いられる。さらに、応答パケット34を用いて肯定応答パケットを送信してトランザクションを終了させることができる。コマンドパケット32と同様に、応答パケット34はさまざまな種類の応答に対するソースノード番号、ソースユニット番号、およびソースタグを含み得る（図5に任意のフィールドとして示す）。

【0029】

図6は、データパケット（dataパケット）36を示す。データパケット36は、図6の実施例において、8ビットリンク上の8つのビット時間を含む。データ

パケット36は、送信されるデータの量に依存して、異なった数のビット時間を含み得る。たとえば、一実施例においてはキャッシュブロックは64バイトを含み、したがって8ビットリンク上の64のビット時間を含む。他の実施例では、キャッシュブロックのサイズを所望により別に定義し得る。さらに、キャッシュ不可能な読出および書込に対しては、キャッシュブロックサイズよりも小さなサイズでデータを送信し得る。キャッシュブロックサイズより小さなデータを送信するためのデータパケットは、より少ないビット時間を用いる。一実施例においては、キャッシュブロックサイズではないデータはデータを送信する前にいくつかのビット時間のバイトイネーブルを送信して、データパケット内でどのデータバイトが有効であるかを示し得る。さらに、キャッシュブロックデータは最初に、要求されたアドレスの最下位のビットによってアドレスされたクワッドワードを備えて返され、その後インターリーブされた残りのクワッドワードのリターンが続く。一実施例においては、クワッドワードは8バイトを含み得る。

【0030】

図3から図6は、8ビットリンクのためのパケットを示す。16および32ビットリンクのためのパケットは、図3から図6に示す連続的なビット時間を連結することにより形成し得る。たとえば、16ビットリンク上のパケットのビット時間1は、8ビットリンク上のビット時間1および2の間に送信される情報を含み得る。同様に、32ビットリンク上のパケットのビット時間1は、8ビットリンク上のビット時間1から4までの間に送信される情報を含み得る。以下の式(1)および式(2)は、8ビットリンクによるビット時間における、16ビットリンクのビット時間1および32ビットリンクのビット時間1の構成を示す。

【0031】

【数1】

$$BT_{16}[15:0] = BT_2[7:0] \parallel BT_1[7:0] \quad (1)$$

$$BT_{32}[31:0] = BT_4[7:0] \parallel BT_3[7:0] \parallel BT_2[7:0] \parallel BT_1[7:0] \quad (2)$$

【0032】

図7は、コンピュータシステム10内のコヒーレントリンクの1つの例示的な実施例によって用いられるコマンドを示すテーブル38を示す。パケットおよびコマンドフィールド符号化のいかなる好適な組をも含む、他の実施例も可能であり企図される。テーブル38は、コマンドの各々に割当てられたコマンド符号化を示すコマンド符号化列、コマンドの名前を示すコマンド列、およびどのコマンドパケット30-34（および特定される場合には、データパケット36）がそのコマンドに対して用いられるかを示すパケットタイプ列を含む。

【0033】

読出トランザクションは、ReadSized, RdBlk, RdBlkSまたはRdBlkModのコマンドのうち、1つを用いて開始される。ReadSizedコマンドは、キャッシュ不可能な読出のために、またはサイズの合ったキャッシュブロック以外のデータの読出のために用いられる。読出されるべきデータ量は、ReadSizedコマンドパケット内に符号化される。キャッシュブロックの読出には、以下の場合以外にRdBlkコマンドを用いることができる。すなわち、(i) キャッシュブロックの書込可能なコピーを所望である場合。この場合はRdBlkModコマンドを用い得る。または(ii) キャッシュブロックのコピーを所望するが、ブロックを変更する意図があるとは分らない場合。この場合はRdBlkS コマンドを用いることができる。RdBlkS コマンドを用いて、ある種のコヒーレントな方式（たとえばディレクトリに基づくコヒーレントな方式）をより効率化できる。一般的に、適切な読出コマンドはソースから送信されて、キャッシュブロックに対応するメモリを有するターゲットノードへの読出トランザクションを開始する。ターゲットノード内のメモリコントローラは、システム内の他のノードに（プローブ応答からのトランザクションのソースへのリターンを示す）Probeコマンドを送信して、これらのノード内のキャッシュブロックの状態を変化させること、およびキャッシュブロックの更新されたコピーを含むノードにキャッシュブロックをソースノードに送らせることにより、コヒーレンスを維持する。Probeコマンドを受取ったノードの各々は、ProbeResp応答パケットをソースノードに送信する。プローブされたノードが読出データの更新されたコピー（すなわちダーティデータ）を有していれば、そのノードはRdResponse応答パケットおよびダーティデータを送信する。ダーティ

データを送信するノードはまた、ターゲットノードによる要求された読出データの送信をキャンセルしようと試みて、ターゲットノードにMemCancel応答パケットをも送信し得る。さらに、ターゲットノード内のメモリコントローラは、RdResponse応答パケットとその後に続くデータパケット内のデータとを用いて要求された読出データを送信する。ソースノードがプローブされたノードからRdResponse応答パケットを受取れば、その読出されたデータが用いられる。そうでなければ、ターゲットノードからのデータが用いられる。一旦ソースノードにおいてプローブ応答および読出されたデータの各々が受取られると、ソースノードはトランザクションの終了の肯定応答として、ターゲットノードにSrcDone応答パケットを送信する。

【0034】

書込トランザクションはWrSizedまたはVicBlkコマンドを用いて開始され、その後に関連のデータパケットが続く。WrSizedコマンドは、キャッシュ不可書込またはサイズの合わないキャッシュブロックのデータの書込に用いられる。WrSizedコマンドに対するコヒーレンシの維持のために、ターゲットノード内のメモリコントローラはシステム内の他のノードの各々に（トランザクションのターゲットノードに対するプローブ応答のリターンを示す）Probeコマンドを送信する。Probeコマンドに応答して、プローブされたノードの各々はターゲットノードにProbeResp応答パケットを送信する。もしプローブされたノードがダーティデータを記憶していれば、プローブされたノードはRdResponse応答パケットおよびダーティデータで応答する。このようにして、WrSizedコマンドによって更新されたキャッシュブロックは、メモリコントローラに返されて、WrSizedコマンドによって提供されるデータとマージされる。メモリコントローラは、プローブされたノードの各々からプローブ応答を受取ると、ソースノードにTgtDoneパケットを送信してトランザクションの終了の肯定応答を提供する。ソースノードはSrcDone応答パケットで応答する。

【0035】

ノードによって変更され、かつノード内のキャッシュで置き換えられたヴィクティムキャッシュブロックは、VicBlkコマンドを用いてメモリに返される。VicB

lkコマンドに対してはプローブは必要ではない。したがって、ターゲットメモリコントローラがビクティムブロックデータをメモリにコミットする準備ができたとき、ターゲットメモリコントローラはビクティムブロックのソースノードにTgtDone応答パケットを送信する。ソースノードは、データがコミットされるべきことを示すSrcDone応答パケットか、またはデータが（たとえば介入するプローブに応答して）VicBlkコマンドの送信とTgtDone応答パケットの受信との間で無効化されたことを示すMemCancel応答パケットのいずれかで応答する。

【0036】

ソースノードに記憶された書込不可能状態のキャッシュブロックへの書込許可を得るために、ソースノードはChangetoDirtyコマンドパケットを送信し得る。ChangetoDirtyコマンドによって開始されるトランザクションは、ターゲットノードがデータを返さないという点を除いて、読出と同様に動作し得る。もしソースノードがキャッシュブロック全体を更新する意図があるのであれば、ValidateBlkコマンドを用いて、ソースノードによって記憶されていないキャッシュブロックへの書込許可を得ることができる。そのようなトランザクションに対してはソースノードへデータは送られないが、それ以外では読出トランザクションと同様に動作する。

【0037】

Nop infoパケットは、たとえばノード間のバッファ空き表示を送信するために用いられる、非動作パケットである。Broadcastコマンドは、ノード間でメッセージをブロードキャストするために用い得る（たとえば、ブロードキャストコマンドは割込を分散させるために用い得る）。最後に、sync infoパケットは、ファブリックの同期が所望である場合（たとえば、エラー検出、リセット、初期化）に用い得る。

【0038】

テーブル38はまた、バーチャルチャネル列（Vchan）をも含む。バーチャルチャネル列は、パケットの各々が移動する（すなわち、パケットの各々が属する）バーチャルチャネルを示す。この実施例においては、3つのバーチャルチャネルが定義される：コマンド（Command、C）、応答（Response、R）、およびプ

ローブ (Probe、P) である。バーチャルチャネルを以下に詳細に説明する。情報パケットは隣接するノード間の通信のために用いられ、よってこの実施例においてはバーチャルチャネルには割当てられないことに留意されたい。

【0039】

バーチャルチャネル

図8を参照すると、バーチャルチャネルを例示するブロック図が示される。図8において、2つのバーチャルチャネルが示される (バーチャルチャネル40Aおよび40B)。処理ノード12A-12Dの各々は、バーチャルチャネル40A-40Bに結合される。2つのバーチャルチャネルは例示目的でのみ図8に示す。他の実施例は、いかなる好適な数のバーチャルチャネルをも含み得る。たとえば、コンピュータシステム10の実施例は以下の図9に示されるように4つのバーチャルチャネルを用い得る。

【0040】

一般的に、「バーチャルチャネル」は、さまざまな処理ノードの間でパケットを担持するための通信経路である。各バーチャルチャネルは、他のバーチャルチャネルから資源的に独立している (すなわち、1つのバーチャルチャネル内で流れるパケットは、物理的な通信の意味において、一般的に別のバーチャルチャネルにおけるパケットの存在または不在によって影響されない)。パケットは、パケットタイプに基づいてバーチャルチャネルに割当てられる。同じバーチャルチャネル内のパケットは、互いの通信に対して物理的にコンフリクトし得る (すなわち、同じバーチャルチャネル内のパケットには資源コンフリクトが生じる可能性がある) が、異なったバーチャルチャネル内のパケットの通信に対しては物理的にコンフリクトしない。

【0041】

ある種のパケットは、他のパケットとロジック的にコンフリクトする (すなわち、プロトコルの理由、コヒーレンシの理由、または他のそのような理由から、一方のパケットはロジック的に別のパケットとコンフリクトし得る)。もし第1のパケットが、ロジック/プロトコル理由から、第2のパケットがそのデスティネーションノードに到着する前にそのデスティネーションノードに到着しなければ

ばならない場合、第2の packets が物理的に第1の packets 通信をブロックするのであれば、(コンフリクトする資源を占有することにより) コンピュータシステムがデッドロックする可能性がある。第1および第2の packets を別々のバーチャルチャネルに割当てることにより、かつ別々のバーチャルチャネルにおける packets が互いの通信をブロックしないようコンピュータシステム内の送信媒体を実現することにより、デッドロックのない動作を実現し得る。異なったバーチャルチャネルからの packets は、同じ物理的リンク (たとえば、図1のライン24) をわたって通信されることに留意されたい。しかしながら、受信バッファは送信の前に利用可能になっているので、バーチャルチャネルはこの共用資源を用いる間にも、互いをブロックしない。

【0042】

1つの観点から見ると、異なった packets タイプの各々 (たとえば、異なったコマンド符号化の各々) は、それ自身のバーチャルチャネルに割当てられることができる。しかしながら、バーチャルチャネルが物理的にコンフリクトしないことを確実にするためのハードウェアは、バーチャルチャネルの数とともに増大する。たとえば、一実施例においては、別々のバッファがバーチャルチャネルごとに割当てられる。別々のバッファがバーチャルチャネルごとに用いられるために、1つのバーチャルチャネルからの packets は別のバーチャルチャネルからの packets と物理的にコンフリクトしない (そのような packets は別のバッファに入れられるためである)。しかしながら、バッファの数はバーチャルチャネルの数に比例する。したがって、ロジック/プロトコル態様でコンフリクトしないさまざまな packets の種類を組合せることにより、バーチャルチャネルの数を減じることが望ましい。そのような packets は、同じバーチャルチャネルを移動する際に互いに対して物理的にコンフリクトし得るが、ロジック的コンフリクトを有さないために、デッドロックなしに資源コンフリクトを解決することを可能にする。同様に、ロジック的にコンフリクトするおそれのある packets を別々のバーチャルチャネルに保つことは、packets 間の資源コンフリクトをなくす。したがって、ロジック的コンフリクトは、第1に完了させるべき packets の進行を可能にすることにより packets 間の資源コンフリクトをなくして、解決される。

【0043】

一実施例においては、コヒーレントリンク上で特定のバーチャルチャネル内で特定のソースノードから特定のデスティネーションノードに移動するパケットは順序を保つ。しかしながら、異なったバーチャルチャネルを特定のソースノードから特定のデスティネーションノードに移動するパケットは、順序付けられない。同様に、（たとえ同じバーチャルチャネルを移動するとしても）特定のソースノードから異なったデスティネーションノードへのパケット、または異なったソースノードから同じデスティネーションノードへのパケットは、順序付けられない。

【0044】

バーチャルチャネルは、コヒーレントファブリックおよび非コヒーレントファブリック上に物理的にマッピングされる（図19を参照）。たとえば、図1に示すコンピュータシステム10の実施例においては、相互接続は各ノード間の単方向リンクを含む。したがって、さまざまなバーチャルチャネルを移動するパケットは、単方向リンクに物理的に送信される。パケットは、ソースとデスティネーションとの間で中間のノードを通して移動し得る。たとえば、ノード12Aからノード12Dを通して移動するパケットは、ノード12Bおよび12Cを通過し得る。異なったバーチャルチャネルを移動するパケットは、コンピュータシステム10によって異なって経路制御される。たとえば、第1のバーチャルチャネル内でノード12Aからノード12Dを移動するパケットは、ノード12Bを通過し得る一方、ノード12Aからノード12Dを第2のバーチャルチャネル内で移動するパケットは、ノード12Cを通過し得る。ノードの各々は、異なったバーチャルチャネル内のパケットの互いに対する物理的コンフリクトを確実になくするための回路を含む。非コヒーレントファブリックにおいては、I/Oノードからのパケットは、そのI/Oノードとホストブリッジ（図19を参照）との間で互いのI/Oノードを通して通過し得る。I/Oノードは、図8に示されるものと同様の態様でバーチャルチャネルに結合され得ることに留意されたい。

【0045】

以下に詳述される特定の実施例においては、制御パケットバッファは各バーチャル

ャルチャネルに割当てられて、そのバーチャルチャネル内を移動する制御パケットをバッファする。別々のデータパケットバッファがまた、データパケットを担持し得る各バーチャルチャネルに割当てられてもよい。制御パケットバッファ（その各エントリは、比較的小さなビット時間を含む）を分けることにより、およびデータパケットバッファ（その各エントリは、キャッシュブロックを保持するために比較的大きな数のビット時間を含む）を分けることにより、バッファ空間を節約する一方で、それでも好適なデータ記憶装置を提供する。データパケットバッファよりも制御パケットバッファが多く実現される（これは、データパケットはすべて対応の制御パケットを有するが、制御パケットはすべてが対応のデータパケットを有するわけではないためである）。バッファ空間を比較的効率的に用いる一方、スループットは高くなり得る。

【0046】

図9は、コンピュータシステムの一実施例に従って規定されるバーチャルチャネルを示すテーブル42である。他の実施例も可能であり企図される。示される実施例に対しては、3つのバーチャルチャネルが規定される。コヒーレントリンクに対するこれらのバーチャルチャネルに属するパケットを図7に示し、非コヒーレントリンクに対するバーチャルチャネルに属するパケットを図20に示す。

【0047】

コマンドバーチャルチャネルはコマンドパケット（および対応のデータパケット）に対して用いられる。一般的に、コマンドはトランザクションを開始し、よってコマンドは付加的なコマンドパケットの通信を生じさせない。さらに、さまざまなコマンドパケットは、デスティネーション（トランザクションのターゲット）に到達するまでこれらの間に順序がないので、互いに対してロジック／プロトコルコンフリクトを有さない。したがって、コマンドパケットは1つのバーチャルチャネル内に含まれることができる。

【0048】

コマンドは、（コヒーレントファブリックにおけるコヒーレンシの維持のための）プローブコマンドパケットの生成および（データを送りトランザクションの肯定応答を提供するための）応答パケットの生成をもたらし得る。したがって、

プローブパケットおよび応答パケットは、コマンドと同じバーチャルチャネルには含まれない（これにより資源コンフリクトおよびロジック的コンフリクトがデッドロックを生じさせることを防ぐ）。さらに、プローブパケットはプローブ応答および読出応答パケットの生成をもたらす得るので、応答パケットとは異なったバーチャルチャネルに入れられる。

【0049】

応答パケットもまた、付加的な応答パケットを生じさせる可能性がある（たとえば、SrcDoneおよびTgtDoneは、互いを生じさせ得る）。したがって、応答パケットは、他の応答パケットと同じバーチャルチャネルに入れられることにより、それらとのロジック的コンフリクトを生じさせる可能性がある。しかしながら、多数の応答バーチャルチャネルを設けることは、付加的なバーチャルチャネルを扱うための資源要求（すなわち、バッファ）が増大するので、不所望であるおそれがある。応答パケットは、直接的または（たとえば、コマンドパケットに応答して生成されたプローブを介する）間接的のいずれであっても、コマンドパケットの結果である。ノード12A-12D（および以下に示すI/Oノード）は、コマンドパケットとのトランザクションを開始する前に、トランザクションの間に生成され得る応答パケット（どのような応答データパケットも含む）を処理するために十分な資源を割当てよう構成される。同様に、プローブコマンドパケットを生成する前に、ノードはプローブ応答パケットを処理するために十分な資源を割当てよう構成される（応答パケットがそのノードに返される場合）。この態様で、すべての応答パケットがデスティネーションノードによって受取られる。したがって、応答パケットは1つの応答バーチャルチャネルにマージされる。応答パケット（および対応のデータパケット）は、応答バーチャルチャネル内を移動する。

【0050】

最後に、プローブコマンドパケットはプローブバーチャルチャネル内を移動する。プローブは、さまざまなメモリ場所のキャッシュ済コピーと、メモリ場所自体との間のコヒーレンスを維持するために用いられる。メモリコントローラによって処理される第1のコマンドパケットに対応するコヒーレンシアクティビティ

は、次のコマンドパケットが処理される前に完了させなければならない可能性がある。たとえば、もしメモリコントローラの待ち行列が同じキャッシュブロックのコマンドで満たされていれば、最初のコマンドが完了するまでメモリコントローラでさらなるコマンドパケットの処理は起こらない。したがって、プローブコマンドパケット（および応答）は、別々のバーチャルチャネルに与えられて、他のバーチャルチャネル内のパケットとの資源コンフリクトが、プローブコマンドパケットをブロックしないことを確実にする。

【0051】

テーブル42はまた、バーチャルチャネルが適用されるコンピュータシステム10におけるリンクの形成を示す（コヒーレントノードの間のコヒーレントリンクおよび非コヒーレントノードの間の非コヒーレントリンク）。非コヒーレントおよびコヒーレントリンクは両方とも、コマンドおよび応答バーチャルチャネルをサポートする。非コヒーレントリンクはコヒーレンスをサポートせず（プローブコマンドパケットはこれを確実にするために用いられる）、したがって、プローブバーチャルチャネルをサポートしない。

【0052】

バーチャルチャネルーコヒーレントファブリック

図10を参照すると、例示的な処理ノード12Aの一実施例のブロック図が示される。他の処理ノード12B-12Dも同様に構成され得る。他の実施例も可能であり企図される。図10の実施例においては、処理ノード12Aはインターフェイスロジック18A、18B、および18Cおよびメモリコントローラ16Aを含む。さらに、処理ノード12Aは、プロセッサコア52、キャッシュ50、パケット処理ロジック58、および任意で第2のプロセッサコア56および第2のキャッシュ54を含む。インターフェイスロジック18A-18Cはパケット処理ロジック58に結合される。プロセッサコア52および56は、キャッシュ50および54にそれぞれ結合される。キャッシュ50および54はパケット処理ロジック58に結合される。パケット処理ロジック58はメモリコントローラ16Aに結合される。

【0053】

一般的に、パケット処理ロジック58は、処理ノード12Aが結合されるリンクで受取られる制御パケットに応答し、キャッシュ50および54および／またはプロセッサコア52および56に応答して制御パケットを生成し、サービスに対してメモリコントローラ16Aによって選択されるトランザクションに応答してプローブコマンドおよび応答パケットを生成し、かつ別のノードへの送信のために、ノード12Aが中間のノードとなるパケットを、インターフェイスロジック18A-18Cの別のものへ経路制御するよう構成される。インターフェイスロジック18A、18B、18Cは、ロジックを含み、パケットを受けて、パケット処理ロジック58によって用いられる内部クロックにパケットを同期させる。

【0054】

パケット処理ロジック58は、コンピュータシステム10によってサポートされるバーチャルチャネルから独立して資源をサポートするためのハードウェアを含む。たとえば、パケット処理ロジック58はバーチャルチャネルごとに別々のバッファを提供する。図11に例示的な実施例を示す。代替的な実施例は、インターフェイスロジック18A-18C内、または他のいずれかの好適な場所にバーチャルチャネルから独立して資源を提供するためのハードウェアを提供する。

【0055】

キャッシュ50および54は、高速キャッシュメモリを含み、該高速キャッシュメモリは、データのキャッシュブロックを記憶するよう構成される。キャッシュ50および54は、プロセッサコア52および56の内部にそれぞれ統合され得る。これに代えて、キャッシュ50および54は所望のようにバックサイドキャッシュ構成またはインライン構成でプロセッサコア52および56に結合されてもよい。さらに、キャッシュ50および54は、キャッシュの階層として実現されてもよい。(階層内で)プロセッサコア52および56に近い方のキャッシュは、所望であればプロセッサコア52および56内に統合されてもよい。

【0056】

プロセッサコア52および56は、予め定められた命令セットに従って命令を実行するための回路を含む。たとえば、x86命令セットアーキテクチャが選択

される。これに代えて、アルファ、パワーPC、または他のどのような命令セットアーキテクチャを選択してもよい。一般的に、プロセッサコアはデータおよび命令のためにキャッシュにアクセスする。もしキャッシュミスが検出されると、読出要求が生成され、ミスのあるキャッシュブロックがマッピングされたノード内のメモリコントローラに送信される。

【0057】

図11を参照すると、パケット処理ロジック58の一実施例のブロック図が示される。他の実施例が可能であり企図される。図11の実施例においては、パケット処理ロジック58は、第1の組の制御およびデータパケットバッファ60、第2の組の制御およびデータパケットバッファ62、第3の組の制御およびデータパケットバッファ64、制御ロジック66、データバッファプール68、および応答カウンタプール70を含む。制御およびデータパケットバッファ60は、コマンドバッファ(CB)60A、応答バッファ(RB)60B、プローブバッファ(PB)60C、コマンドデータバッファ(CDB)60D、および応答データバッファ(RDB)60Eを含む。同様に、制御およびデータパケットバッファ62は、コマンドバッファ(CB)62A、応答バッファ(RB)62B、プローブバッファ(PB)62C、コマンドデータバッファ(CDB)62D、および応答データバッファ(RDB)62Eを含む。制御およびデータパケットバッファ64は、コマンドバッファ(CB)64A、応答バッファ(RB)64B、プローブバッファ(PB)64D、コマンドデータバッファ(CDB)64Dおよび応答データバッファ(RDB)64Eを含む。制御およびデータパケットバッファ60はインターフェイスロジック18Aによって受取られたパケットを受取るよう結合される(たとえば、ライン24B上)。同様に、制御およびデータパケットバッファ62は、インターフェイスロジック18Bによって受取られたパケットを受取るよう結合され、制御およびデータパケットバッファ64はインターフェイスロジック18Cによって受取られたパケットを受取るよう結合される。制御およびデータパケットバッファ60、62、および64は、制御ロジック66に結合される。さらに、応答データバッファ60E、62E、および64Eが、データバッファプール68に結合される。データバッファプール68

および応答カウンタプール70は、制御ロジック66に結合され、これはさらにノードIDレジスタ72、制御パケットアクティブレジスタ74A-74Cおよびデータパケットアクティブレジスタ76A-76Cを含む。制御ロジック66は、受信および送信インタフェースを介してインターフェイス18A-18Cに結合され、かつ同様にメモリコントローラ16Aおよびキャッシュ50（および任意のキャッシュ54）に結合される。データバッファプール68がさらにメモリコントローラ16Aおよびキャッシュ50（および任意のキャッシュ54）に結合される。

【0058】

制御およびデータパケットバッファの組の各々が、バーチャルチャネルごとに異なったバッファを提供する。すなわち、本実施例においては、コマンドバッファ60Aはコマンドバーチャルチャネルに割当てられ、応答バッファ60Bは応答バーチャルチャネルに割当てられ、プローブバッファ60Eはプローブバーチャルチャネルに割当てられる。この態様で、1つのバーチャルチャネルにおける制御パケットの受信は、別のバーチャルチャネルにおける制御パケットの受信によって妨げられない。各バーチャルチャネルからの制御パケットは、そのバーチャルチャネルに対応する制御パケットバッファに記憶され、よって別のバーチャルチャネルからの制御パケット（異なった制御パケットバッファに記憶される）と物理的コンフリクトを起こさない。同様に名称がつけられるバッファ62および64内のバッファは、上述のようにバーチャルチャネルに割当てられる。

【0059】

同様に、データパケットバッファは、データパケットを担持する各バーチャルチャネルに与えられる（プローブバーチャルチャネルはこの実施例においてはデータパケットを担持しない）。すなわち、本実施例においては、コマンドデータバッファ60Dはコマンドバーチャルチャネルに割当てられ、応答データバッファ60Eは応答バーチャルチャネルに割当てられる。同様に名称がつけられるバッファ62および64内のバッファは、上述のようにバーチャルチャネルに割当てられる。

【0060】

本実施例においては、インターフェイスロジック18A-18Cは、受取ったパケットを制御パケット（制御経路内に提供される）と、データパケット（データ経路に提供される）とに分割するよう構成される。制御経路は、制御パケットバッファに結合され（たとえば、バッファ60A-60Cがインターフェイスロジック18Aからの制御経路に結合される）、データ経路はデータパケットバッファに結合される（たとえば、バッファ60D-60Eがインターフェイスロジック18Aからのデータ経路に結合される）。制御ロジック66は、受信および送信インターフェイスを介してパケットの表示を受取るよう結合され、かつ受取ったパケットのバッファエントリを割当てよう構成される。他の企図される実施例においては、受取ったパケットはインターフェイスロジックによって制御およびデータパケットに分割されない。そのような実施例においては、制御ロジック66はCTL信号を受取って、ビット時間のデータパケットとビット時間の制御パケットとを区別する。

【0061】

一般的に、制御ロジック66は、他のバッファ内のパケットから独立してさまざまなバッファからのパケットを処理するよう構成される。したがって、異なるバーチャルチャネル内を移動するパケットは互いに対して物理的にコンフリクトしない。

【0062】

制御ロジック66は、バッファ60、62および64内の制御パケットを調べて、制御パケットがノード12A（「このノード」）に対して宛先決めされているか、または別のノードに送られるべきかどうかを判断する。ノードIDレジスタ72は、このノードのノードIDを記憶し、制御ロジック66はノードIDを用いて制御パケットがこのノードに対して宛先決めされているかどうかを判断する。本実施例においては、プローブバーチャルチャネル内のパケットはブロードキャストパケットであり、よってこのノードと、このノードがパケットを送るべき他のノードに対して宛先決めされる（よってノードID比較は用いられない）。他のバーチャルチャネル内のパケットは方向付けられた（directed）パケットであり、これに対してデスティネーションノードフィールドは、パケットがこの

ノードに宛先決めされているか、または別のノードに送られるべきであるかどうかを識別する。制御ロジック66は、1つ以上の経路制御テーブルを含み、これはデスティネーションノードごとに、インターフェイスロジック18A-18Cのどれがパケットを送るために用いられるべきかを示す。制御ロジック66は、識別されたインターフェイスロジック18A-18Cを介して送信された受信パケットを受けよう結合される受信ノードが、制御パケットに対応するバーチャルチャネルに対する空いた制御パケットバッファを有している場合に、制御パケットを送り得る。さらに、制御パケットがデータパケットを識別すると、制御パケットに対応するバーチャルチャネルに対する空いたデータパケットバッファは、制御パケットおよび後続の特定されたデータパケットを送る前に識別される。制御ロジック66は、制御パケット（およびもし特定されていればデータパケット）が送られるべきであるかどうかを判断し、受信および送信インターフェイスを用いてパケットを識別されたインターフェイスロジック18A-18Cに送り、これは次いでパケットを受信ノードに送る。また、制御ロジック66は、対応の種類のバッファが空いたことを記憶して、次の情報パケットがインターフェイス18A-18Cを介して送られるようにし、パケットはノード12Aによって受取られて、受信インターフェイスの送信ノードに対する空きバッファを示す。

【0063】

一方、もし制御パケットがこのノードに宛先決めされていれば、制御ロジック66はパケットの種類に依存してパケットを処理する。たとえば、もし制御パケットがメモリコントローラ16Aに標的を定められたコマンドであれば、制御ロジック66は制御パケットをメモリコントローラ16Aに運ぶことを試みる。メモリコントローラ16Aは、たとえば処理されるべきトランザクションに対して待ち行列を用い、かつ待ち行列がフルであれば制御パケットを拒絶し得る。プローブパケットを処理するために、制御ロジック66はキャッシュ50および54（およびプロセッサコア52および56の内部のいずれのキャッシュでも）と通信して、アドレス指定されたキャッシュブロックのステータスを判断する。制御ロジック66は、ステータスを備えたプローブ応答パケット（またはもしキャッシュブロックがノード内で変更されていれば、読出応答パケットとデータ）を生

成し、プローブ応答パケットを送信する（受信ノードバッファが利用可能である場合）。

【0064】

受取ったパケットの処理に加えて、制御ロジック66は、キャッシュ50および54からのフィル要求およびヴィクティムブロックに応答してパケットを生成し、かつプロセッサコア52および56からの直接の要求に応答してパケットを生成する（たとえば、キャッシュ不可能要求、I/O要求など）。さらに、応答パケットはメモリコントローラに応答して生成され、送信に対するデータを提供、またはトランザクションを完了させる。制御ロジック66は、メモリコントローラ16に応答して、処理に対する対応のコマンドを選択してプローブコマンドパケットを生成し、プローブコマンドパケットをブロードキャストし得る（他のパケット送信と同様に、受信ノードバッファが利用可能であると想定する）。

【0065】

上述のように、ノードはそのノードによって生成された制御パケットに対応する応答パケットを処理するための十分な資源を提供する。本実施例においては、制御ロジック66はパケットを生成するが、このパケットは2つの場合においてノードに返される応答パケットをもたらし得る：（i）トランザクションを開始するためのコマンドパケットを生成する場合（たとえば、キャッシュ50および54、またはプロセッサコア52および56からの要求に応答する）；および（ii）メモリコントローラ16Aを標的とする制御パケットに対するプローブパケットを生成する場合。より特定的には、（ii）の場合は、メモリコントローラ16Aを標的とするサイズ決めされた書込の場合に生じ得る。いずれの場合においても、制御ロジック66は、応答パケットを処理するための資源を割当ててゐる。

【0066】

本実施例においては、制御ロジック66は応答を処理するためにデータバッファプール68および応答カウンタプール70からの資源を割当ててゐる。データバッファプール68は、データのキャッシュブロックを記憶するための複数のエントリを含み得る一方、応答カウンタプール70は複数のカウンタを含み得る。データバッファプールエントリは、トランザクションに対応する応答データを記憶す

るために割当てられ得る。カウンタは、受取った応答をカウントする（かつ、プローブ応答において提供され得るいかなる状態情報をも保持する）ために割当てられ得る。応答パケットは、割当てられたカウンタを用いて（期待される数の応答に到達するまで）カウントされ、応答パケットによって受取られたデータは割当てられたデータバッファに記憶され得る。トランザクションに關与する最大2つのパケットがデータを担持し得ることに留意されたい（1つは、MemCancel応答パケットが、応答パケットの送信の前にメモリコントローラに到達しなかった場合の、標的決めされたメモリコントローラからのものであり、他方はデータの変更されたキャッシュ済コピーを有するプローブされたノードからのものである）。2つのデータパケットが受取られた場合には、プローブされたノードからのパケットが保持されて、メモリコントローラからのパケットは捨てられる。

【0067】

一旦期待される応答の各々が受取られ、応答データが受取られていると、制御ロジック66は行なわれているトランザクションのタイプに依存して、データをメモリコントローラ16Aまたはキャッシュ50または54に、送信する。たとえば、もし応答がパケット処理ロジック58によって生成されたプローブコマンドに応答して生成されたプローブ応答であれば、応答データはメモリコントローラ16Aに送信され得る。これに代えて、もし応答が読出トランザクションによるものであれば、データはキャッシュ50または54に送信される。

【0068】

データバッファプール68は、ノード12Aから送信されるべきデータを記憶するためにも用いられることに留意されたい。たとえば、ノード12Aからの書込コマンドに対するヴィクティムブロックデータまたは書込データは、データバッファプール68に記憶され得る。これに代えて、別々のバッファがこのデータのために提供されてもよい。さらに、さまざまなトランザクションに対して用いられるバッファのプールを提供する代わりに、所望のように、トランザクションタイプによって別々のバッファが提供されてもよい。

【0069】

ここで用いられる場合、バッファとは後の検索のために情報の1つ以上の項目

を記憶するために用いられる記憶要素である。バッファは、1つ以上のレジスタ、ラッチ、フリップフロップ、または他のクロックされる記憶装置を含む。これに代えて、バッファは好適に構成された組のランダムアクセスメモリ（RAM）セルを含んでもよい。バッファは、いくつかのエントリに分割され、各エントリはバッファが設計されている情報の1つの項目を記憶するよう設計される。エントリは、いかなる好適な様式においても割当てられおよび割当て解除され得る。たとえば、バッファはより古いエントリが削除されるにつれエントリがシフトされる、シフト先入れ先出し（FIFO）バッファで動作されてもよい。これに代えて、ヘッドおよびテールポインタを用いて、バッファ内の最も古いおよび最も新しいエントリを示し、エントリはバッファから削除されるまでバッファの特定の記憶場所に保持されてもよい。ここで使用される「制御ロジック」という用語は、説明された動作を実現するために、入力に対して動作し、これに応答して出力を生成する、組合わせロジックおよび／またはステートマシンのいかなる組合せをも指す。

【0070】

特定の一実施例においては、パケットは一連のビット時間としてインターフェイスロジック18A-18Bから受取られる。インターフェイスロジック18A-18Cは、制御またはデータビット時間が送信されているかどうかを示し、制御ロジック66は、ビット時間を記憶するための適切なバッファを生じさせる。制御ロジック66は、制御パケットアクティブレジスタ74およびデータパケットアクティブレジスタ76を用いて、現在受取られている制御パケットまたはデータパケットがどのバーチャルチャネルに属しているかを識別し得る。制御パケットアクティブレジスタ74は、インタフェースロジック18A-18Cごとに与えられる（たとえば、制御パケットアクティブレジスタ74Aはインターフェイス18Aに対応し得る）。同様に、データパケットアクティブレジスタ76はインターフェイスロジック18A-18Cごとに与えられる（たとえば、データパケットアクティブレジスタ76Aはインターフェイス18Aに対応し得る）。第1のビット時間の制御パケットに応答して、制御ロジック66はコマンドフィールド（第1のビット時間内）をデコードし、制御パケットがどのバーチャルチ

チャネルに割当てられているかを判断する。制御ロジック66は、バッファエントリを対応の（制御パケットが受取られるインターフェイスロジック18A-18Cに対応する組内の）制御パケットバッファに割当て、および、パケットが受取られるインターフェイスロジック18A-18Cに対応する制御パケットアクティブレジスタ76が、その制御パケットバッファを示すよう設定する。同じインターフェイスロジック18A-18Cからの後続の制御パケットビット時間は、各ビット時間の制御パケットが受取られるまで、示されたバッファ内の示されたエントリに記憶される。もし制御パケットがデータパケットを特定すると、制御ロジック66は、データパケットバッファエントリを識別されたバーチャルチャネルに対応するデータパケットバッファに割当てる。データパケットビット時間は、各ビット時間のデータが受取られるまで、示されたバッファの示されたエントリ内に記憶される。代替的な実施例においては、インターフェイスロジック18A-18Cは、ビット時間のパケットを集め、パケット全体をパケット処理ロジック58に送信してもよい。そのような実施例においては、制御パケットアクティブレジスタ74およびデータパケットアクティブレジスタをなくしてもよい。さらに別の実施例においては、インターフェイスロジック18A-18Cは、パケット処理ロジック58への並行した送信のためのいくつかのビット時間を集めるが、ビット時間の数がパケットよりも少ない可能性がある。さらに別の実施例においては、バッファ60、62および64はパケット処理ロジック58の代わりにそれぞれのインターフェイスロジック18A-18C内に場所決めされてもよい。

【0071】

図11に示される実施例においては、インターフェイスロジック18A-18Cごとにバッファの別々の組が提供される。代替的な実施例においては、バッファはインターフェイスロジックの間で分割され得る（バーチャルチャネルタイプごとの）プールとして提供されてもよい。そのような実施例は、別のノードには結合されないインターフェイスロジック（たとえば、図1の例におけるインターフェイスロジック18C）に、ゼロバッファを与えることによりバッファを活用し得る。インターフェイスロジック18Cに割当てられたであろうバッファは、

インターフェイスロジック18A-18Bによって用いられるために割当てられる。

【0072】

図12を参照すると、データバッファプール68内のデータバッファプールエントリ80の一実施例を例示する図が示される。他の実施例も可能であり企図される。図12の実施例においては、データバッファプールエントリ80は、ソーススタグフィールド82、ソースノードフィールド84、ソースユニットフィールド88、およびデータフィールド86を含む。

【0073】

制御ロジック66が、トランザクションに対する応答データパケットを記憶するためにデータバッファプールエントリ80を割当ててる場合、制御ロジック66はトランザクションのソースノード、ソースユニット、およびソーススタグを、ソースノードフィールド84、ソースユニットフィールド88、およびソーススタグフィールド82にそれぞれ記憶する。ソースノード、ソースユニット、およびソーススタグは独自に未処理の(outstanding)トランザクションを識別し、かつソースノード、ソースユニット、およびソーススタグは未処理のトランザクションに対応する応答パケットによって担持されるために、トランザクションの応答パケット(および対応のデータパケット)が識別されて、データパケットは割当てられたエントリに記憶され得る。言換えると、応答データパケットを特定する応答パケットが受取られると、応答パケットのソースノード、ソースユニット、およびソーススタグは、ソースノードフィールド84、ソースユニットフィールド88およびソーススタグフィールド84と比較されて、応答データのために先行して割当てられたデータバッファプールエントリ80の位置を突止め、データは応答データバッファからデータバッファプールエントリ80のデータフィールド86にコピーされ得る。データフィールド86はデータのキャッシュブロックを含み得る。

【0074】

図13を参照すると、応答カウンタプール70内の応答カウンタ90の一実施例を例示する図が示される。他の実施例も可能であり企図される。図13の実施

例においては、応答カウンタ90はソースタグフィールド92、ソースノードフィールド94、ソースユニットフィールド95、応答カウントフィールド96、および受信ステートフィールド98を含む。

【0075】

制御ロジック66が、トランザクションに対する応答カウントを記憶するために応答カウンタ90を割当てする場合、制御ロジック66はソースノードフィールド94、ソースユニットフィールド95、およびソースタグフィールド92にそれぞれトランザクションのソースノード、ソースユニットおよびソースタグを記憶し得る。ソースノードフィールド94、ソースユニットフィールド95、およびソースタグフィールド92は、データバッファプールエントリ80の対応のフィールド84、88、および82と同様に用い得る。

【0076】

トランザクションの割当の際に、応答カウントフィールド96は、トランザクションに対して期待される応答の数に初期化され得る。ソースノード、ソースユニット、およびソースタグがフィールド94、95および92にそれぞれ記憶された応答パケットが受取られると、応答カウントがデクリメントされる。応答カウンタが0に到達すると、すべての応答は受取られ、トランザクションはコミットされている。これに代えて、カウントは0に初期化されて、応答パケットは期待される数の応答が受取られるまで応答カウントをインクリメントしてもよい。

【0077】

応答ステートフィールド98は、データが受取られるであろう状態を示すために用い得る。状態は、ノード12Aがキャッシュブロックの受信の際に獲得した、キャッシュブロックに対するアクセス権およびキャッシュブロックに対するコヒーレンシを維持するための責任を示す。たとえば、MOESI（変更、所有、排他的、共用、および無効）コヒーレンシ状態が用いられて、受信状態フィールド98は、サポートされる状態のうちの1つに符号化される。これに代えて、他のいかなる好適な組のコヒーレンシ状態が用いられてもよい（たとえばMESI状態）。受信状態フィールド98は、どの他のノードもトランザクションによって送られるキャッシュブロックのコピーを有さない場合に獲得される状態に初期

化され得る。プローブ応答が受取られて、キャッシュブロックのコピーがプローブノードによって維持されているか、またはダーティデータが提供されていることを示すと、受信状態フィールド98はそれに従って更新される。一実施例においては、共用ビットがプローブ応答パケットに含まれて、キャッシュブロックのコピーがプローブ応答を提供したプローブされたノードによって維持されていることを示し得る。さらに、プローブされたノードからの読出応答パケットの受信は、ノードがキャッシュブロックのダーティコピーを有することを示す。読出応答パケットはまた、プローブされたノードによってコピーが維持されているかどうかを示すための共用ビットを含み得る。

【0078】

データバッファプール68および応答カウンタプール70は、未処理のトランザクションに対する応答を扱うための一例にすぎないことに留意されたい。別の実施例においては、未処理のトランザクションのテーブルが維持される。テーブルは、上述のものと同様のソースノード、ソースユニット、ソースタグ、データ、受信状態、および応答カウント（または、制御ロジック66がすべての応答が受取られたことを判断し得る等価の情報）を含む。いかなる好適な組の資源をも用い得る。

【0079】

図14は、パケットを受信するためのパケット処理ロジック58の一実施例の動作を示すフローチャートである。他の実施例も可能であり企図される。示される実施例は、パケットを一連のビット時間としてパケット処理ロジック58内に受取る。他の実施例は、ビット時間のパケットをインターフェースロジック18A-18Cに累積させ、完全なパケットをパケット処理ロジック58に提供するが、この場合パケットの受信をビット時間で管理することに関連するステップは削除される。図14に示されるステップは、理解を助けるために特定の順序で示されているが、いかなる好適な順序をも用い得る。さらに、ステップはパケット処理ロジック58内の組合せロジックを用いて並列に行なわれてもよい。図14に示されるステップは、インターフェースロジック18A-18Cの各々から独立して並列に行われ得るが、これはインターフェースロジックごとにビット時間が

並行に受信されるためである。

【0080】

パケット処理ロジック58は、受信したビット時間がデータパケットの一部であるか、またはコマンドパケットの一部であることを示すインターフェ이스ロジックからの信号を受取る。もしビット時間がデータパケットビット時間であれば（判断ブロック100）、ビット時間はそのインターフェ이스ロジックに対応するデータパケットアクティブレジスタによって示されるデータバッファ（およびデータバッファ内のエントリ）によって記憶される（ステップ102）。もしデータパケットビット時間がデータパケットの最後のビット時間であると、制御ロジック66は対応のデータパケットアクティブレジスタを無効化する。一方、もしビット時間が制御パケットビット時間であれば、パケット処理ロジック58は制御パケットが現在受取りが進行中のものであるかどうかを判断する（たとえば、制御パケットアクティブレジスタが有効であるかどうか、判断ブロック104）。もし制御パケットが現在進行中であれば、ビット時間は制御パケットアクティブレジスタによって示される制御パケットバッファに記憶される（ステップ106）。もし制御パケットビット時間が制御パケットの最後のビット時間であれば、制御ロジック66は対応の制御パケットアクティブレジスタを無効化し得る。

【0081】

これに代えて、制御パケットは現在進行中ではない可能性がある。この場合には、パケット処理ロジック58は新しく受取られた制御パケットのコマンドフィールドをデコードして、制御パケットが属するバーチャルチャネルを識別する（ステップ108）。識別されたバーチャルチャネルに対応する制御パケットバッファエントリは割当てられ、制御パケットビット時間は割当てられた制御パケットバッファエントリに記憶される。

【0082】

さらに、パケット処理ロジック58は、制御パケットが後続のデータパケットを特定するかどうかを判断する（判断ブロック110）。データパケットが特定されると、パケット処理ロジック58は識別されたバーチャルチャネルに対応するデータバッファからデータバッファエントリを割当て、そのデータバッファ（

およびエントリ)を示すようデータパケットアクティブレジスタを更新する(ステップ112)。

【0083】

ここで図15を参照して、コマンドパケットを処理するためのパケット処理ロジック58の一実施例の動作を示すフローチャートを示す。他の実施例も可能であり意図される。図15に示されるステップは理解を容易にするためある特定の順序にて示されるが、任意の好適な順序を用いてよい。さらに、ステップは、並列に、組合せロジックをパケット処理ロジック58内にて用いながら実行してもよい。図15に示されるステップは並列かつ独立して各インターフェイスロジック18A~18Cおよび/または各コマンドパケットバッファごとに実行されてもよく、なぜならば、異なるインターフェイスおよび/または異なるバーチャルチャネルからのコマンドパケットは物理的に独立しているからである。代替的に、1つのコマンドパケット(またはインターフェイスロジック18A~18Cにつき1つのコマンドパケット)を、処理に対し、好適な公平性アルゴリズムに従って選択してもよい。一般に、1つのバーチャルチャネルから処理のために選択されたパケットはあるチャネル内におけるパケットに対する順序付け規則に従う(たとえば、同じソースから同じデスティネーションへのパケットは順に選択される)が、パケットの、処理に対する選択は、所望されかつ順序付け規則が許す場合には、順序が乱れた状態でなされてもよい。

【0084】

パケット処理ロジック58はコマンドパケットの標的がこのノードであるかどうかを判断する(判断ブロック126)。たとえば、パケット処理ロジック58はコマンドパケットのデスティネーションノードフィールドに記録されるデスティネーションノードをノードIDレジスタ72に記憶されるノードIDと比較してもよい。これらのノードが一致する場合、そのコマンドはこのノードを標的としている。コマンドがこのノードを標的としない場合には、パケット処理ロジック58はそのコマンドパケット(および特定される場合には対応のデータパケット)をデスティネーションノードに応答して送る(ステップ128)。たとえば、パケット処理ロジック58は、パケットを特定のデスティネーションノードへ

送るためインターフェースロジック18A~18Cの1つを送信インターフェースとして識別するパケット経路制御テーブルを維持してもよい。対応のコマンドバッファ（およびデータパケットが特定される場合にはデータバッファ）が、パケット経路制御テーブルにより特定されるリンクに結合される受信ノードにおいて利用可能であるなら、パケット処理ロジック58はコマンドパケットを送る。加えて、コマンドパケットがデータパケットを特定する場合には、コマンドパケットを送ることは、送信リンク上のデータパケットが活性状態ではあるが未だ送信されていない場合には、ストールされてもよい。

【0085】

コマンドパケットがこのノードを標的としている場合、パケット処理ロジック58はコマンドパケット（および適用可能である場合には対応のデータパケット）をメモリコントローラ16Aに与えてもよい（ステップ130）。一旦コマンドパケットが処理される（送られるかまたはこのノードによって受け入れられる）と、そのコマンドパケットはコマンドバッファエントリから取除かれ（適用可能である場合にはデータはコマンドデータバッファエントリから取除かれ）ることに注目されたい。

【0086】

プローブコマンドは同じ態様で処理されてもよいが、ただし、プローブコマンドは後のデータパケットを特定せず、したがって、データパケットに対するチェックは無視されるかもしれないことに注意されたい。さらに、プローブは内部的に（たとえば当該ノード内においてキャッシュをプローブすることにより）処理されるとともに送られてもよく、なぜならば、それらは放送パケットであるからである。ノードはキャッシュのプローブ後にプローブ応答パケットを発生し送信してもよい。

【0087】

選択されるコマンドパケットが対応のデータパケットを特定する場合には、さまざまな実施例によってそのコマンドパケットの処理をたとえそのデータパケットがまだ受取られていなくても行なってもよく、またはそのデータパケットの到着を待つことによりそのデータを送ることを単純化するかまたは完全なデータパ

ケットを特定する別の制御パケットが同じリンク上において送られるようにしてもよい。コマンドパケットが処理されるときにデータパケットが受取られていなかった場合には、そのデータパケットはそれが受取られるときに上述のように処理されてもよい。

【0088】

図16を参照して、応答パケットを処理するためのパケット処理ロジック58の一実施例の動作を示すフローチャートを示す。他の実施例も可能であり意図される。図16に示されるステップは理解を容易にするためある特定の順序で示されるが、任意の好適な順序を用いてよい。さらに、ステップは、並列に、組合せロジックをパケット処理ロジック58内で用いながら実行してもよい。図16に示されるステップは並列かつ独立して各インターフェースロジック18A~18Cおよび／または各応答パケットバッファごとに実行してもよく、なぜならば、異なるインターフェースおよび／または異なるバーチャルチャネルからのコマンドパケットは物理的に独立しているからである。

【0089】

パケット処理ロジック58は応答パケットのデスティネーションノードがこのノードであるか否かを判断する（判断ブロック144）。デスティネーションノードが別のノードである場合に、応答バーチャルチャネルに対する空いているバッファエントリがその応答パケットが送られるリンク上の受信機にある場合には、パケット処理ロジック58はその応答パケット（および適用可能である場合には対応のデータパケット）を送る（ステップ146）。

【0090】

応答パケットのデスティネーションノードがこのノードである場合には、パケット処理ロジック58は対応の応答カウンタを減分し受信状態を更新するよう構成されるが、それは、その応答が、受信状態がデフォルト状態から変更されるべきである旨を示すプローブ応答である場合である（ステップ148）。加えて、その応答パケットがデータパケットを特定する場合には、そのデータパケットを、対応の応答データバッファから、その応答に割当てられたデータバッファに移動させる（ステップ150）。

【0091】

カウンタを減分した後、パケット処理ロジックはカウンタをテストしてすべての応答パケットが受取られ処理されたか否かを判断する（判断ブロック152）。すべての応答パケットが受取られ処理されたと判断される場合には、パケット処理ロジック58はメモリコントローラ16Aまたはキャッシュ50および54に対しそれらがコマンドを完了してもよい旨を知らせ、データバッファからの関連のデータおよび応答カウンタからの受信状態を与えてもよい（適用可能である場合—ステップ145）。一旦応答パケットが処理される（送られるかまたはこのノードによって受入れられる）と、その応答パケットは応答バッファエントリ（および適用可能である場合には応答データバッファエントリ）から取除かれる。

【0092】

上の議論において、コマンドパケットまたは応答パケットの処理の「中断」という語が用いられている。一般に、処理が「中断」されるのは、その特定のパケットの処理が、中断の理由が取除かれるまで停止される場合である。同じタイプの他のパケットの処理をそのコマンドまたは応答の処理の中断中に行なってもよい。

【0093】

選択される応答パケットが対応のデータパケットを特定する場合、さまざまな実施例による応答パケットの処理をたとえそのデータパケットがまだ受取られていなくても（つまりそのデータパケットがまだデータバッファ内になくても）行なってもよく、またはデータパケットの到着を待つことによりそのデータを送ることを単純化するかまたは完全なデータパケットを特定する別の制御パケットが同じリンク上において送られるようにしてもよい。応答パケットが処理されるときにデータパケットが到着していない場合には、そのデータパケットはそれが受取られるときに上記のように処理されてもよい。

【0094】

図17を参照して、ノードが結合されるリンク上においてパケットを起動するためのパケット処理ロジック58の一実施例の動作を示すフローチャートを示す

。他の実施例も可能であり意図される。図17に示されるステップは理解を容易にするためある特定の順序にて示されるが、任意の好適な順序を用いてよい。さらに、ステップを、並列に、組合せロジックをパケット処理ロジック58内にて用いながら実行してもよい。パケット処理ロジック58はリンク上のパケットを応答的に起動することによりキャッシュ50および54からの要求／犠牲ブロックならびに／またはコア52および56により実行される動作を満たしてもよい。加えて、プローブパケットの起動を、処理のためにメモリ動作を選択するメモリコントローラ16Aに応答して行なってもよい。応答パケットの起動は、プローブが処理された後、かつこのノードによりソースを与えられるかまたはこのノードを標的としているトランザクションが完了されたことに応答して行なわれてもよい。

【0095】

パケット処理ロジック58は、起動されるべきパケットの結果、データがこのノードに戻されるかもしれないか否かを判断する（判断ブロック160）。たとえば、ノードによって起動される読出トランザクションはデータがノードに戻されることを引き起こし、一方、ノードにより起動される書込トランザクションはデータがノードに戻されることを引き起こさない。CleanToDirtyトランザクションの結果、データはノードに戻されるかもしれない（別のノードがキャッシュブロックに対しダーティ状態において影響を与えた場合）。同様に、プローブパケットはデータをこのノードへ戻すことを引き起こすかもしれないが、それは、別のノードがキャッシュブロックに対しダーティ状態において影響を及ぼしプローブ応答がこのノードに向けられる場合である。トランザクションの結果、データがこのノードに戻される場合には、パケット処理ロジック58はあるデータバッファをデータバッファプール68から割当て（ステップ162）。

【0096】

加えて、パケット処理ロジック58はプローブ応答がこのノードに対しパケットに応答して戻されるか否かを判断する（ステップ166）。これが生ずるのはパケットがプローブであるかまたはパケットがこのノードへのプローブ応答をもたらすようなトランザクションを開始している場合であるかもしれない（たと

ば読出トランザクション)。プローブ応答がこのノードに戻される場合、パケット処理ロジック58は応答カウンタを応答に対し当該トランザクションに割当てその応答カウンタをコヒーレントファブリックにおけるノードの数に初期化する(ステップ168)。

【0097】

パケット処理ロジック58は、さらに、他の応答が、このノード(たとえばSrcDone、TgtDoneなど)に対し、起動中のパケットに応答して戻されるかどうかを判断する(ステップ164)。そのような他の応答が戻される場合には、パケット処理ロジック58は応答カウンタを割当て初期カウントを1に設定する(ステップ165)。この後、パケット処理ロジック58はパケットを送る(ステップ170)。

【0098】

予め資源を割当てることにより応答パケット(データを含む)の取扱いをトランザクションの起動前に行なうことにより、応答パケットは受信で処理可能である。したがって、ある応答パケットは他の応答パケットとロジック的/プロトコルコンフリクトを有するかもしれないが、応答パケットは応答バーチャルチャネルにマージされるかもしれないが、なぜならば、物理的コンフリクトは各応答パケットがそのデスティネーションノードに到着するにつれ処理されることにより取除かれるかもしれないからである。

【0099】

次に図18を参照して、バッファ解放フィールドを含む情報パケット180の一実施例を示すブロック図が示される。他の実施例も可能であり意図される。バッファ解放フィールドは各バッファのタイプごとに含まれる。RespDataフィールドは応答データバッファに対応する。Responseフィールドは応答バッファに対応する。同様に、CmdDataフィールドはコマンドデータバッファに対応し、Cmdフィールドはコマンドバッファに対応する。Probeフィールドはプローブバッファに対応する。

【0100】

バッファ解放フィールドの各々は2つのビットを含み、単一のリンク上におけ

る送信機から受信機への1つの情報パケット180の伝送において解放されるべき0～3の各タイプのバッファを斟酌する。3つより多いエントリをバッファ内において与えてもよく、複数の情報パケットを用いて1つのタイプのうちの3つより多くを解放してもよい。パケット処理ロジック58はバッファカウントを各タイプのバッファおよび各インターフェ이스ロジック18A～18Cごとに含んで、各インターフェイスが結合されるリンクの他方端における受信機によって与えられる各タイプのバッファの総数を示してもよい。これらのカウンタの初期化は、パワーアップにおいて、情報パケットを受信機から送信機へ、バッファ解放フィールドがその受信機において利用可能なバッファ数にセットされる状態で送ることにより行なわれてもよい。3つより多いエントリは、複数の情報パケットを送ることにより示されてもよい。

【0101】

パケット処理ロジック58はパケットをある所与のバーチャルチャネルにおいて送ってもよいが、それは、対応のタイプのバッファ（およびパケットがデータパケットを特定する場合にはデータバッファ）が、パケットが送られている受信機において利用可能である限りにおいてである。加えて、パケット処理ロジック58は、ノード12Aにおいてパケット処理ロジック58によるパケットの処理のため解放される、各インターフェイス18A～18Cに対する各タイプのバッファの数に注目する。周期的に、パケット処理ロジック58は情報パケット180を各インターフェ이스ロジック18A～18Cを介して送り、そのリンク上の送信機に対し、パケット処理ロジック58により解放されたバッファエントリの数を示す。

【0102】

バーチャルチャネルー非コヒーレントファブリック

ここで図19を参照して、I/Oサブシステム200の一実施例のブロック図を示す。他の実施例も可能であり意図される。図19の実施例においては、I/Oサブシステム200はホストブリッジ202と複数のI/Oノード204A、204Bおよび204Cとを含む。ホストブリッジ202は、処理ノード12Dに、ライン24I～24Jを含むコヒーレントリンクを介して結合され、さらに

は、I/Oノード204Aに、ライン24K~24Lを含む非コヒーレントリンクを用いて結合される。I/Oノード204A~204Cはさらなる非コヒーレントリンクを用いてデージーチェーン構成で相互接続される（ライン24N~24O）。

【0103】

一般に、I/Oノード204A~204CはトランザクションをI/Oサブシステム200内にて起動してもよい。それらトランザクションは究極的には別のI/Oノード204A~204C、別の非コヒーレントリンク上のI/Oノード、またはメモリ14にて標的としてもよい。単純化のため、トランザクションはホストブリッジ202とI/Oノード204A~204Cとの間においてその実際の標的にかかわらず実行されてよい。ホストブリッジ202はトランザクションをI/Oサブシステム202内において処理ノード12A~12Dからの要求に代わって起動してもよく、当該コヒーレントファブリックまたはコンピュータシステム内の他のホストブリッジを標的としているI/Oノード204A~204Cにより起動されるトランザクションを取扱ってもよい。したがって、I/Oノード204A~204Cにより送られるパケットはホストブリッジ202に向かってデージーチェーン接続を介して流れてもよい（「上流に」流れる）。ホストブリッジ202により送られるパケットは受信側I/Oノード204A~204Nの方向に流れてもよい（「下流に」流れる）。I/Oノードとホストブリッジとをデージーチェーンで相互接続しI/Oノードが（トランザクションレベルにおいて）ホストブリッジとのみ通信することにより、I/Oノードが他のノードではなくホストブリッジに直接接続されるように見えるI/Oサブシステム200のロジック的な眺めが与えられる。

【0104】

I/Oサブシステム200はホストブリッジに対しデージーチェーン相互接続の両端で接続されることによりリンク欠陥の場合における堅牢性を可能にするかまたは共有されるI/Oサブシステムを処理ノードのクラスタ間にて可能にしてもよい。1つのブリッジはマスタブリッジとして定義され、他のブリッジはスレーブブリッジとして規定されるであろう。リンク欠陥がない場合、サブシステム

内におけるすべてのI/Oノードはマスタブリッジに属してもよい。リンク欠陥の検出で、その欠陥の両側のノードは再プログラミングされることによりホストブリッジに対しその欠陥の同じ側で属し、それによって、2つの異なるサブシステムを形成し処理ノードとの通信を維持する。I/Oノードを、2つのホストブリッジ間において、欠陥がない状態で分配する（2つのロジック的に別々のチェーンを形成する）ことにより、トラフィックも同様にバランスをとってもよい。

【0105】

パケットがデージーチェーンの端部（たとえば図19の例においてはI/Oノード204C）に到達しI/Oノード204A～204Cのいずれもそのパケットを受入れない場合、エラーをそのチェーンの端部にあるノードによって発生してもよい。

【0106】

一般に、I/Oサブシステム200はリンクを非コヒーレント相互接続として実現してもよい。その非コヒーレントリンクにおけるデータパケット定義は図6に示され記載されるものと同様であってもよく、その非コヒーレントリンクにおける情報パケット定義は図3および図18において（Probeフィールドが予約される状態で）示されるパケット定義と同様であってもよい。コマンドおよび応答パケットは下に図21および図22において示される。

【0107】

バーチャルチャネルに関し、非コヒーレントリンクは上記のコヒーレントリンクと同じバーチャルチャネルを用いてもよい。しかしながら、プローブは非コヒーレントリンクにおいては用いられないので、プローブバーチャルチャネルは取除かれてもよい。図9に示される表42は非コヒーレントリンクの一実施例に対して定義されるバーチャルチャネルを示す。

【0108】

ホストノード202は処理ノード12から分離して示されているが、ホストノード202は所望される場合には処理ノードに統合されてもよいことに注目されたい。

【0109】

次に図20を参照して、テーブル210はコンピュータシステム10における非コヒーレントリンクの一例示的实施例に従って用いられるパケットを示す。他の実施例も可能であり意図され、任意の他の好適な組のパケットおよびコマンドフィールドエンコーディングを含む。テーブル210は、各コマンドに割り当てられるコマンドエンコーディングを示すコマンドコードコラムと、非コヒーレントパケットが属するバーチャルチャネルを定義するバーチャルチャネル (Vchan) コラムと、コマンドを名づけるコマンドコラムと、コマンドパケット30、212および214のうちどれがそのコマンドに対して用いられるかを示すパケットタイプコラムとを含む。

【0110】

Nop、WrSized、ReadSized、RdResponse、TgtDone、Broadcast、およびSyncパケットは、図7に関して記載された対応のコヒーレントパケットと同様であってもよい。しかしながら、非コヒーレントリンクでは、プローブパケットもプローブ応答パケットも出されない。

【0111】

AssignおよびAssignAckパケットを用いてUnit IDをノードに割り当てる。マスタホストブリッジはパケットを各ノードに（一度に1つ）割り当て、最後に用いられたUnit IDを示す。受信ノードは、そのノードによって必要とされるUnit IDの数を、最後に用いられたUnit ID+1から開始して割り当てる。受信ノードは、割り当てられたUnit IDの数を示すIDカウントを含んで、AssignAckパケットを戻す。

【0112】

次に図21を参照して、非コヒーレントリンクにおいて用いられてもよいコマンドパケット212の一実施例のブロック図を示す。コマンドパケット212はコヒーレントパケットと同様なコマンドフィールドを含む。加えて、随意のソースタグフィールドを、コヒーレントコマンドパケットと同様に、ビット時間3において含んでもよい (SrcTag)。アドレスはビット時間5-8（および最下位アドレスビットに対しては随意にビット時間4）に含まれる。しかしながら、ソースノードの代わりに、ユニットIDが与えられる。

【0113】

ユニットIDはパケットを同じロジックソースから来るものとして認識するよう働く（ユニットIDが等しい場合）。しかしながら、I/Oノードは複数のユニットIDを有してもよい（たとえばノードがロジック的に分離している複数のデバイスまたは機能を含む場合）。したがって、ノードは1つを超えるユニットIDを有するパケットを受入れてもよい。加えて、パケットはホストブリッジとノードとの間を流れるので、パケットに関与するノードの1つ（ホストブリッジノード）が当該パケットに対して暗示される。したがって、単一のユニットIDを非コヒーレントパケットにおいて用いてもよい。一実施例では、ユニットIDは5つのビットを含んでもよい。ユニットID0はホストブリッジに割当てられてもよく、ユニットID31はエラーケースに対して用いられてもよい。したがって、30までのユニットIDが1つのデージーチェーン内へ結合されるI/Oノードにおいて存在してもよい。

【0114】

加えて、コマンドパケット212はシーケンスIDフィールドをビット時間1および2に含む。このシーケンスIDフィールドを用いて同じユニットIDから2つ以上のコマンドパケットの組をグループ化しその組が順序づけられる旨を示してもよい。より特定のには、シーケンスIDフィールドが0である場合、パケットは順序づけられない。シーケンスIDフィールドが0でない場合、パケットは同じシーケンスIDフィールド値を有する他のパケットに関し順序づけられる。

【0115】

次に図22を参照して、非コヒーレントリンクにおいて用いられてもよい応答パケット214の一実施例のブロック図が示される。応答パケット214は、コマンドフィールドと、ユニットIDフィールドと、コマンドパケット212に類似するソースタグフィールドとを含む。他のビットを所望のように含んでもよい。

【0116】

ここで図23を参照して、I/Oノード204Aの一実施例を示すブロック図が示される。他のI/Oノード204B～204Cも同様に構成されてもよい。

他の実施例も可能であり意図される。図23の実施例では、I/Oノード204 Aは、インターフェースロジック18Mおよび18Nと、パケットバッファの第1の組220と、パケットバッファの第2の組222と、ノードロジック224とを含む。インターフェースロジック18Mはライン24Kおよび24Lと、パケットバッファ220およびノードロジック224とに結合される。インターフェースロジック18Nは、パケットバッファ222およびノードロジック224と並んで、ライン24Mおよび24Nに結合される。ノードロジック224はさらにパケットバッファ222および224に結合される。

【0117】

インターフェースロジック18Mおよび18Nはパケットをライン24Lおよび24Mから（それぞれ）受取りパケットをライン24Kおよび24Nで（それぞれ）送るよう構成される。上記においてコヒーレントリンクに対し記載されたインターフェースロジックと同様に、インターフェースロジック18Mおよび18Nは受取られたパケットを制御経路およびデータ経路に分離してもよい。制御経路は制御パケットバッファに結合され、データ経路はデータパケットバッファに結合される。代替的に、インターフェースロジックは受取られたパケットを制御経路およびデータ経路に分離せず、ノードロジック224によって各ビット時間に対応するCDL信号を受取り分離を実行してもよい。コヒーレントインターフェイスと同様に、パケットバッファ220および222は、コマンドバッファ（CB）および応答バッファ（RT）を制御パケットに対して、非コヒーレントリンクにおいて実施される2つのバーチャルチャネルに対応して含む。加えて、データパケットバッファが各バーチャルチャネルに対して与えられる（つまりコマンドデータバッファ（CDB）および応答データバッファ（RDB））。

【0118】

ノードロジック224はバッファ222および224内に受取られたパケットを処理し、パケットをI/Oノード204Aにより実現される周辺機能性に応答して起動してもよい。図11に示される制御ロジック66と同様に、ノードロジック224は制御パケットアクティブレジスタ226Aおよび226B（パケットバッファ222および224にそれぞれ対応する）ならびにデータパケットア

クティブレジスタ228Aおよび228B（パケットバッファ222および224にそれぞれ対応する）を実現してもよい。加えて、非コヒーレントリンクはNode IDの代わりにUnit IDに従って動作するので、ノードロジック224は1つ以上のUnit IDレジスタ230A～230Nを含んでI/Oノード204Aに割当てられるUnit IDを記憶してもよい。Unit IDレジスタ230A～230Nの数は、ノードによって、当該ノード内にて実現されるUnit IDの数に従って変動してもよい。

【0119】

異なるバーチャルチャネルにおけるパケットはI/Oノード204A内の異なるバッファに記憶されるので、異なるバーチャルチャネルにおけるパケットは物理的には互いとコンフリクトしない。したがって、デッドロックのない動作が与えられてもよい。加えて、ノードロジック224は予め資源を割当てて資源パケットおよび応答データを取扱ってもよく（上述のとおり）、したがって応答パケットは1つのバーチャルチャネル内にマージされてもよい（上述のとおり）。ノードロジック224は、さらに、図18に示されるパケットと同様のNopパケットの送受（予約されたプローブフィールドとともに）をフロー制御バッファ220および222（および他のノードにおける同様のバッファ）に対してパケットの送受に関し行なうよう構成されてもよい。

【0120】

ノードロジック224は、さらに、I/Oノード204Aにより実行されるさまざまなI/Oまたは周辺機能に対応するロジックを含んでもよい。たとえば、I/Oノード204Aは記憶周辺装置、たとえばディスクドライブ、CD ROM、DVDドライブなどを含んでもよい。I/Oノード204Aは通信周辺装置、たとえばIEEE 1394、イーサネット（R）、ユニバーサル・シリアル・バス（USB）、周辺構成要素相互接続（PCI）バス、モデムなどを含んでもよい。任意の好適なI/O機能をI/Oノード204Aに含んでよい。

【0121】

ここで図24を参照して、パケットを受信するためのノードロジック224の一実施例の動作を示すフローチャートを示す。他の実施例も可能であり意図され

る。示される実施例はパケットをバッファ220および222に一連のビット時間として受取る。他の実施例はパケットのビット時間をインターフェイスロジック18M~18Nにおいて累積し、完全なパケットをバッファ220および222に与えてもよく、その場合には、ビット時間においてパケットの受取りを管理することに関連するステップは除いてもよい。図24に示されるステップは理解を容易にするためある特定の順序で示されるが、任意の好適な順序を用いてよい。加えて、ステップは、並列に、組合せロジックをノードロジック224内において用いながら実行されてもよい。図24に示されるステップは並列かつ独立して各インターフェイスロジック18M~18Nごとに実行されてもよく、なぜならば、ビット時間は同時に各インターフェイスロジックから受取られてもよいからである。ステップ100~112は上記の図14の対応して記載されたステップと同様であってもよい。

【0122】

次に図25を参照して、コマンドパケットを処理するためのノードロジック224の一実施例の動作を示すフローチャートを示す。他の実施例も可能であり意図されてよい。図25に示されるステップは理解を容易にするためある特定の順序で示されるが、任意の好適な順序を用いてよい。加えて、ステップは、並列に、組合せロジックをノードロジック224内にて用いながら実行されてもよい。図25に示されるステップは並列かつ独立して各インターフェイスロジック18M~18Nごとおよび／または各コマンドパケットバッファごとに実行されてもよく、なぜならば、異なるインターフェイスおよび／または異なるバーチャルチャネルからのコマンドパケットは物理的に独立しているからである。代替的に、1つのコマンドパケット（またはインターフェイスロジック18M~18Nにつき1つのコマンドパケット）を、処理のために、好適な公平性アルゴリズムに従って選択してもよい。一般に、1つのバーチャルチャネルから処理のために選択されるパケットはバーチャルチャネル内におけるパケットに対する順序付け規則に従う（たとえば同じソースから同じデスティネーションへのパケットは順序よく選択される）が、パケットの、処理に対する選択は、所望されかつ順序付け規則が許す場合には、順序が乱れた状態でなされてもよい。

【0123】

ステップ126は一般に図15の対応のステップと同様であってもよい（が、ユニットIDおよびユニットIDレジスタ230A～230Nに基づく）。加えて、ノードロジック224はパケットをパケット経路制御テーブルに従ってではなく同じ方向（上流または下流）において送るよう構成される（ステップ242）。パケットがこのノードを標的としている場合、ノードロジック224はそのパケットを受入れ（パケットを下流のフローから取除く）、それを処理する（ステップ240）。一旦コマンドパケットが処理される（送られるかまたはこのノードによって受入れられる）と、そのコマンドパケットはコマンドバッファエントリから取除かれ（適用可能である場合にはデータパケットはデータバッファエントリから取除かれ）ることに注目されたい。

【0124】

選択されたコマンドパケットが対応するデータパケットを特定する場合には、さまざまな実施例によるコマンドパケットの処理をたとえそのデータパケットがまだ受取られていなくても行なってもよく、またはデータパケットの到着を待つことによりデータを送ることを単純化するかまたは完全なデータパケットを特定する他の制御パケットが同じリンク上において送られるようにしてもよい。データパケットがコマンドパケットの処理時に受取られていない場合、そのデータパケットは、上述のように、それが受取られるときに処理されてもよい。

【0125】

次に図26を参照して、応答パケットを処理するためのノードロジック224の一実施例の動作を示すフローチャートを示す。他の実施例も可能であり意図されてよい。図26に示されるステップは理解を容易にするためある特定の順序で示されるが、他の好適な順序を用いてよい。加えて、ステップは、並列に、組合せロジックをノードロジック224内にて用いながら実行されてもよい。図26に示されるステップは並列かつ独立して各インターフェイスロジック18M～18Nおよび／または各応答パケットバッファごとに実行されてもよく、なぜならば、異なるインターフェイスおよび／または異なるバーチャルチャネルからのコマンドパケットは物理的に独立しているからである。ステップ114は図16の

対応するステップと同様であってもよい（が、ノードロジック224に対してはユニットIDおよびユニットIDレジスタ230A~230Nに基づく）。

【0126】

デスティネーションノードが別のノードである場合、応答バーチャルチャネルに対する空いているバッファエントリが受信機において応答パッケージが送られるリンク上にあるならば、ノードロジック224は応答パッケージ（および適用可能である場合には対応のデータパッケージ）を送る（ステップ250）。この受信機は応答パッケージをそのパッケージが既に流れているのと同じ方向（上流または下流）において流れさせるノードである。

【0127】

応答パッケージのデスティネーションノードがこのノードである場合、ノードロジック224は、もし対応のデータパッケージがあれば、それを、対応の応答データバッファから、その応答が割当てられるデータバッファへ移動させるよう構成される（ステップ252）。換言すれば、ノードロジック224はそのデータを消費する。ノードロジック224は次いで対応のコマンドを完了し、データバッファの割当てを解除する（ステップ254）。一旦応答パッケージが処理される（送られるかまたはこのノードによって受入れられる）と、その応答パッケージは応答バッファエントリから取除かれ（適用可能である場合にはデータパッケージはデータバッファエントリから取除かれ）ることに注目されたい。

【0128】

選択された応答パッケージが対応するデータパッケージを特定する場合には、さまざまな実施例によるその応答パッケージの処理がたとえそのデータパッケージがまだ受取られていなくとも行なってもよく、またはそのデータパッケージの到着を待つことによりそのデータを送ることを単純化するかまたは完全なデータパッケージを特定する別の制御パッケージが同じリンク上において送られるようにしてもよい。データパッケージが応答パッケージの処理時に受取られていない場合、データパッケージは、上述のように、そのデータパッケージが受取られるときに扱ってもよい。

【0129】

次に図27を参照して、ノードが結合されるリンク上においてパッケージを起動

するためのノードロジックの一実施例の動作を示すフローチャートを示す。他の実施例も可能であり意図されてよい。図27に示されるステップは理解を容易にするためある特定の順序で示されるが、他の好適な順序を用いてよい。加えて、ステップは、並列に、組合せロジックをノードロジック224内にて用いながら実行してもよい。

【0130】

ノードロジック224は、開始されるべきトランザクションの結果、データがこのノードに戻されるかどうかを判断する（判断ブロック260）。たとえば、ノードにより開始される読出トランザクションはデータをノードに戻し、一方、ノードにより開始される書込トランザクションはデータをノードに戻さない。トランザクションの結果、データがこのノードに戻される場合、ノードロジック224はデータバッファを割当てて、戻されたデータを記憶する（ステップ262）。この後、ノードロジック224はパケットを送る（ステップ164）。

【0131】

産業上の利用可能性

この発明はコンピュータシステムにおいて適用可能であってよい。

【図面の簡単な説明】

【図1】 コンピュータシステムの一実施例のブロック図である。

【図2】 図1に示される2つのノードの、間のリンクの一実施例が強調される、一実施例のブロック図である。

【図3】 情報パケットの一実施例を示すブロック図である。

【図4】 コヒーレントリンクのためのコマンドパケットの一実施例を示すブロック図である。

【図5】 コヒーレントリンクのための応答パケットの一実施例を示すブロック図である。

【図6】 データパケットの一実施例を示すブロック図である。

【図7】 コヒーレントリンクのためのパケット定義の一実施例を示すテーブルである。

【図8】 バーチャルチャネルの対を示すブロック図である。

【図9】 バーチャルチャネルの組の一実施例を示すテーブルである。

【図10】 図1に示される処理ノードの一実施例のブロック図である。

【図11】 図10に示されるパケット処理ロジックの一実施例のブロック図である。

【図12】 データバッファプールエントリの一実施例を示すブロック図である。

【図13】 応答カウンタプールエントリの一実施例を示すブロック図である。

【図14】 パケット受信のために図10に示されるパケット処理ロジックの一実施例の動作を示すフローチャートである。

【図15】 コマンドパケットを処理するための図10に示されるパケット処理ロジックの一実施例の動作を示すフローチャートである。

【図16】 応答パケットを処理するための図10に示されるパケット処理ロジックの一実施例の動作を示すフローチャートである。

【図17】 パケットを起動するための図10に示されるパケット処理ロジックの一実施例の動作を示すフローチャートである。

【図18】 バッファ空きフィールドを含む情報パケットの一実施例を示すブロック図である。

【図19】 図1および図2に示される相互接続に類似したリンクを介して相互接続されるホストブリッジおよび複数のI/Oノードを含むI/Oサブシステムの一実施例のブロック図である。

【図20】 非コヒーレントリンクのパケット定義の一実施例を示すテーブルである。

【図21】 非コヒーレントリンクのためのコマンドパケットの一実施例のブロック図である。

【図22】 非コヒーレントリンクのための応答パケットの一実施例のブロック図である。

【図23】 I/Oノードの一実施例のブロック図である。

【図24】 パケット受信のための図23に示されるノードロジックの一実

施例の動作を示すフローチャートである。

【図25】 コマンドパケットを処理するための図24に示されるノードロジックの一実施例の動作を示すフローチャートである。

【図26】 応答パケットを処理するための図24に示されるノードロジックの一実施例の動作を示すフローチャートである。

【図27】 パケットを起動するための図27に示されるノードロジックの一実施例の動作を示すフローチャートである。

【図1】

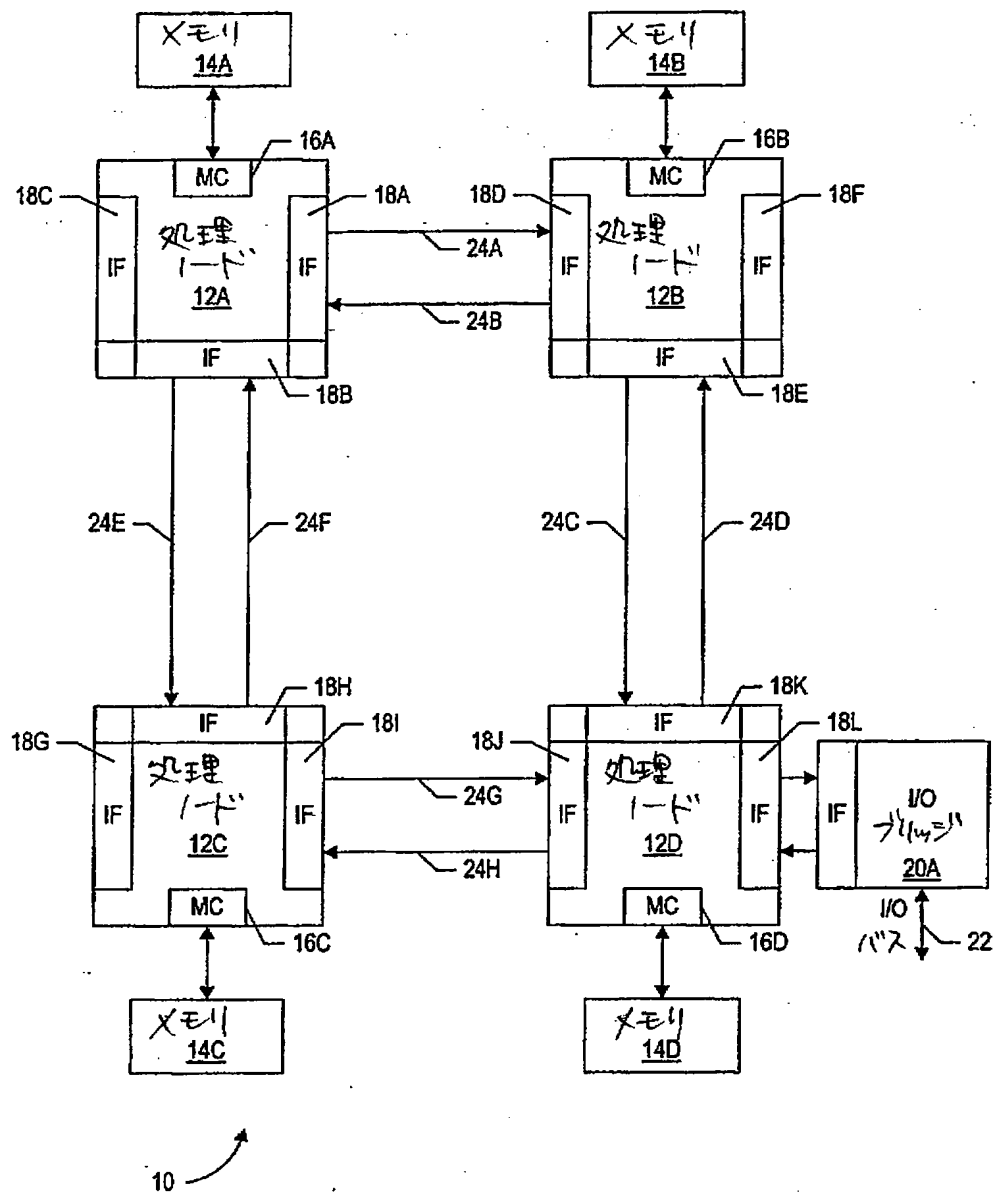


Fig. 1

【図2】

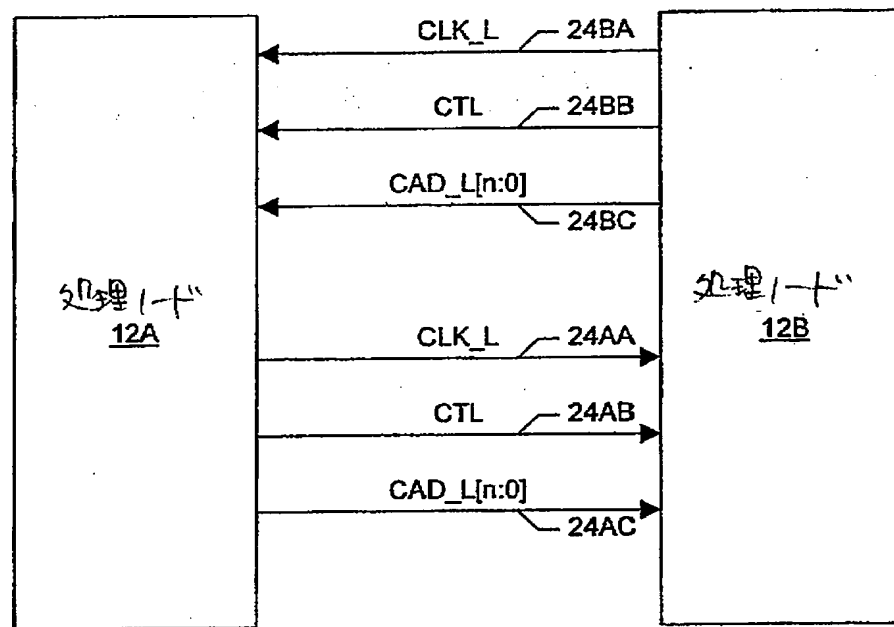


Fig. 2

【図3】

命令時間	7	6	5	4	3	2	1	0
1	CMD[5:0]							
2								
3								
4								

30

Fig. 3

【図4】

ビット時間	7	6	5	4	3	2	1	0
1	SrcUnit [1:0]		CMD[5:0]					
2	DestNode [2:0]			DestUnit [1:0]		SrcNode[2:0]		
3				SrcTag[4:0]				
4								
5	Addr[15:8]							
6	Addr[23:16]							
7	Addr[31:24]							
8	Addr[39:32]							

32

Fig. 4

【図5】

ビット時間	7	6	5	4	3	2	1	0
1	SrcUnit [1:0]		CMD[5:0]					
2	DestNode [2:0]			DestUnit [1:0]		SrcNode[2:0]		
3				SrcTag[4:0]				
4	Sh							


34 

Fig. 5

【図6】

ビット時間	7	6	5	4	3	2	1	0
1	Data[7:0]							
2	Data[15:8]							
3	Data[23:16]							
4	Data[31:24]							
5	Data[39:32]							
6	Data[47:40]							
7	Data[55:48]							
8	Data[63:56]							

36

Fig. 6

【図7】

CMD ¹⁷ 17 ⁹	VChan	コマンド	17 ⁹ 17 ⁹
000000	-	Nop	Info
000001	C	VicBlk	Command/Data
000010		Reserved	
000011	C	ValidateBlk	Command
000100	C	RdBlk	Command
000101	C	RdBlkS	Command
000110	C	RdBlkMod	Command
000111	C	ChangetoDirty	Command
x01xxx	C	WrSized	Command/Data
01xxxx	C	ReadSized	Command
100xxx		Reserved	
110000	R	RdResponse	Response/Data
110001	R	ProbeResp	Response
110011	R	TgtDone	Response
110100	R	SrcDone	Response
110101	R	MemCancel	Response
11011x		Reserved	
11100x	P	Probe	Command
111010	P	Broadcast	Command
111011		Reserved	
11110x		Reserved	
111110		Reserved	
111111	-	Sync	Info

【図8】

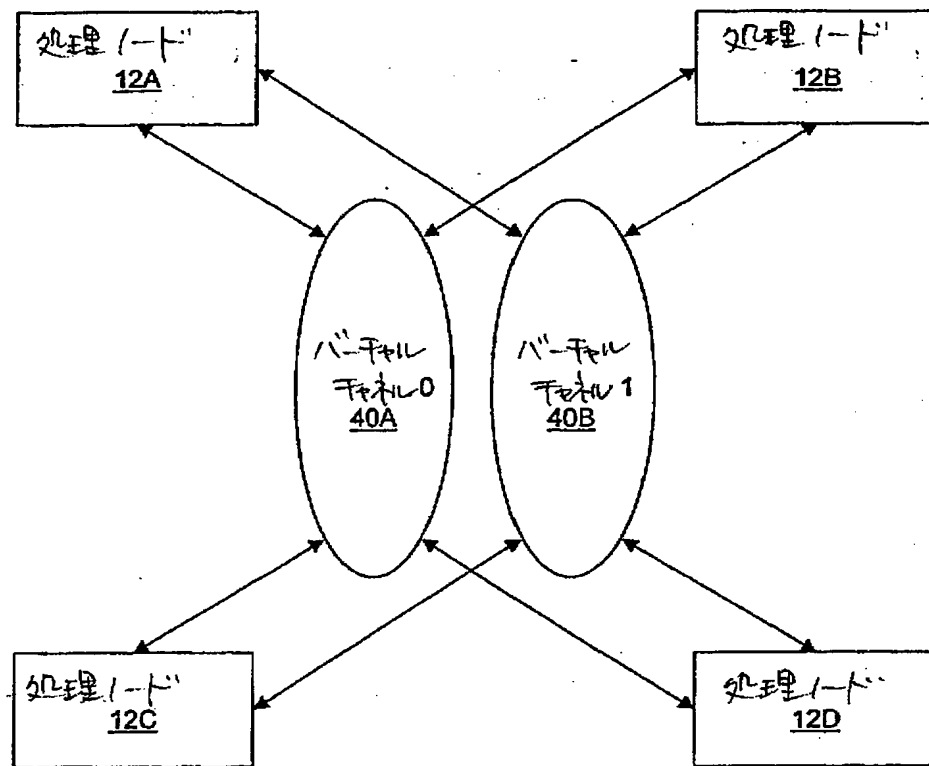


Fig. 8

【図9】

バーチャルチャネル

バーチャルチャネル	適用可能リンク
コマンド	コヒーレントおよび非コヒーレント
応答	コヒーレントおよび非コヒーレント
フロップ	コヒーレントのみ

42

Fig. 9

【図10】

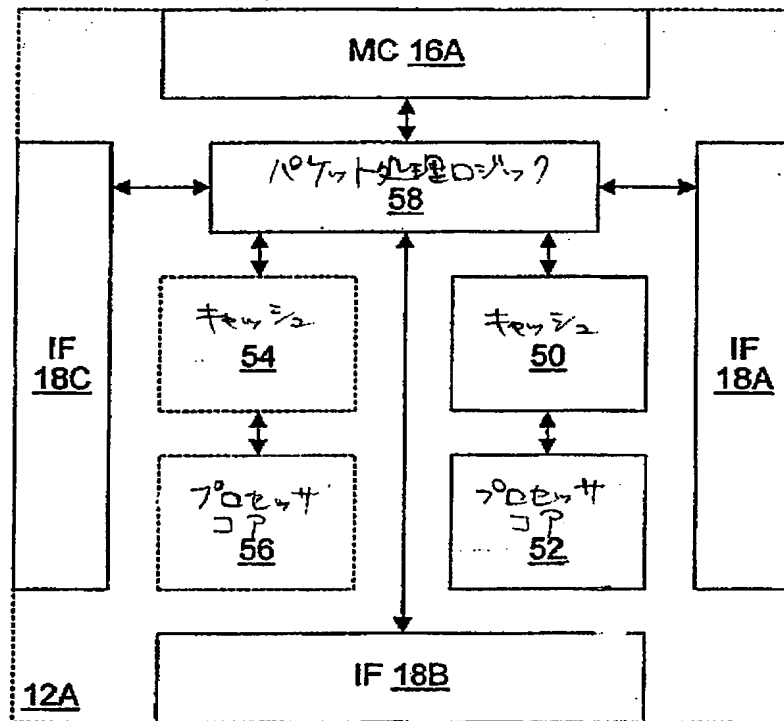
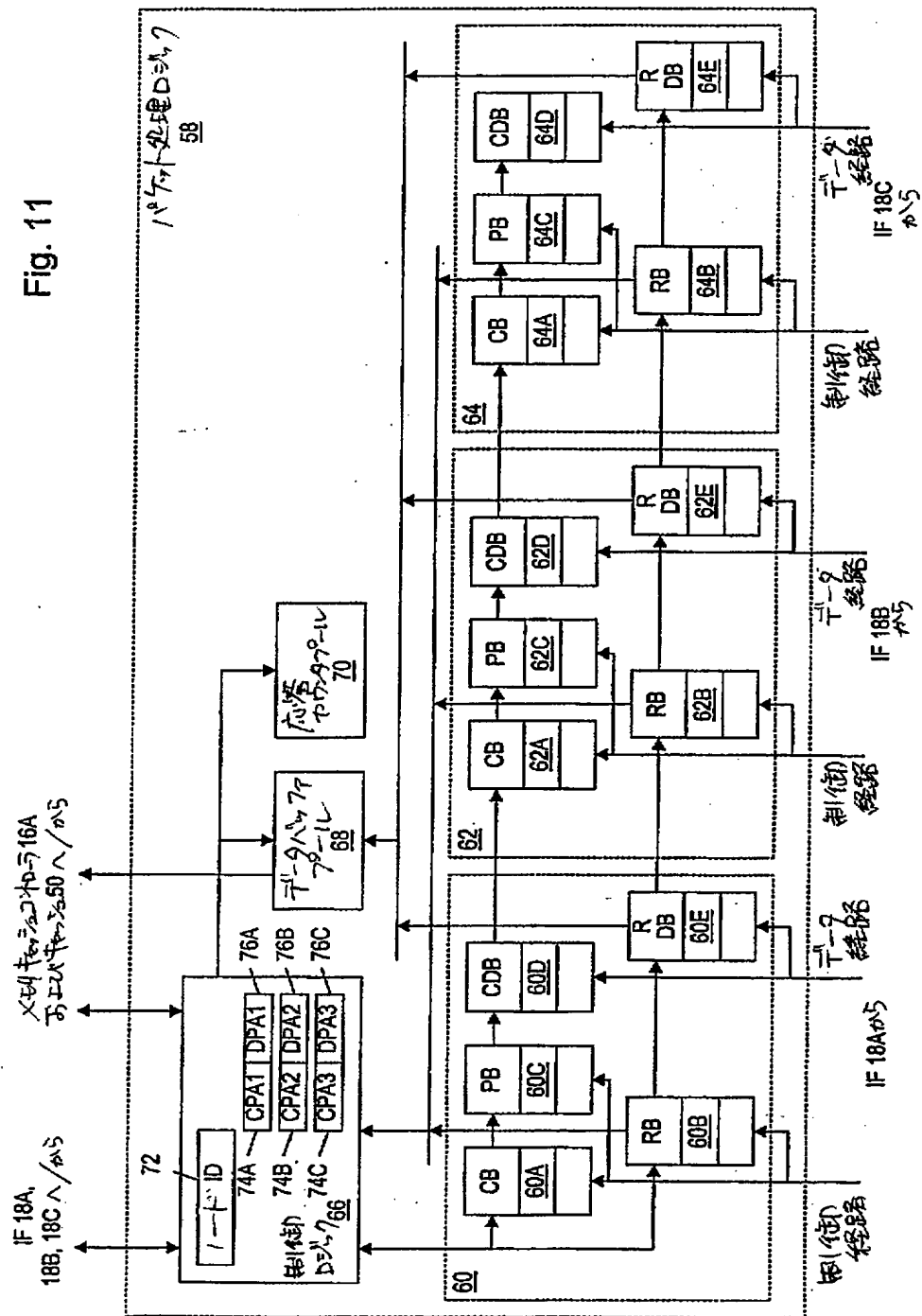


Fig. 10

【図 1 1】

Fig. 11



【図12】

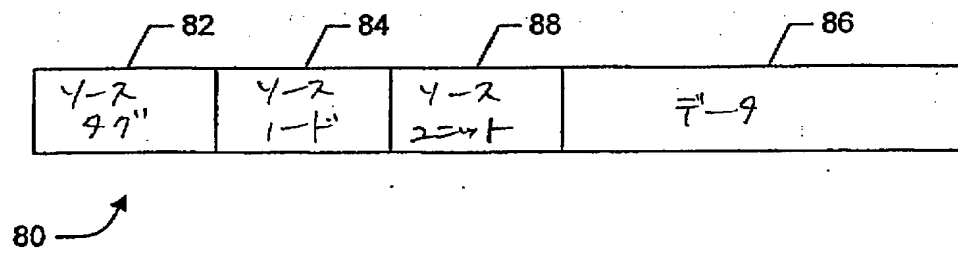


Fig. 12

【図13】

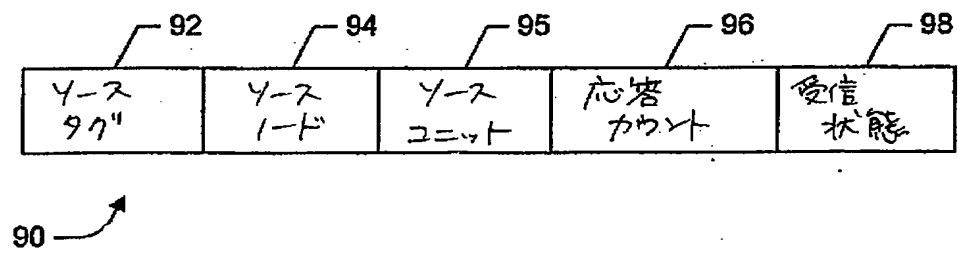


Fig. 13

【図14】

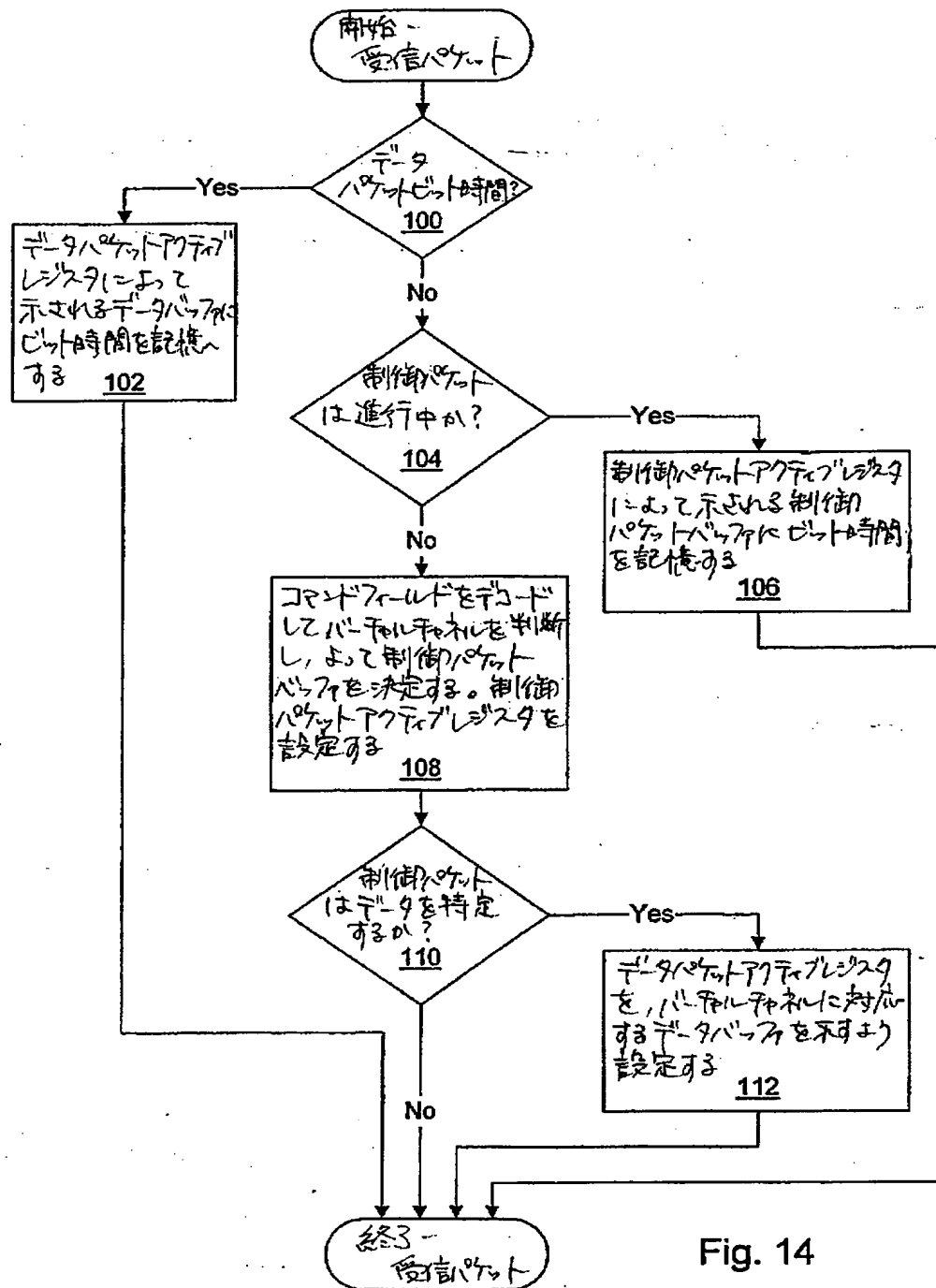


Fig. 14

【図15】

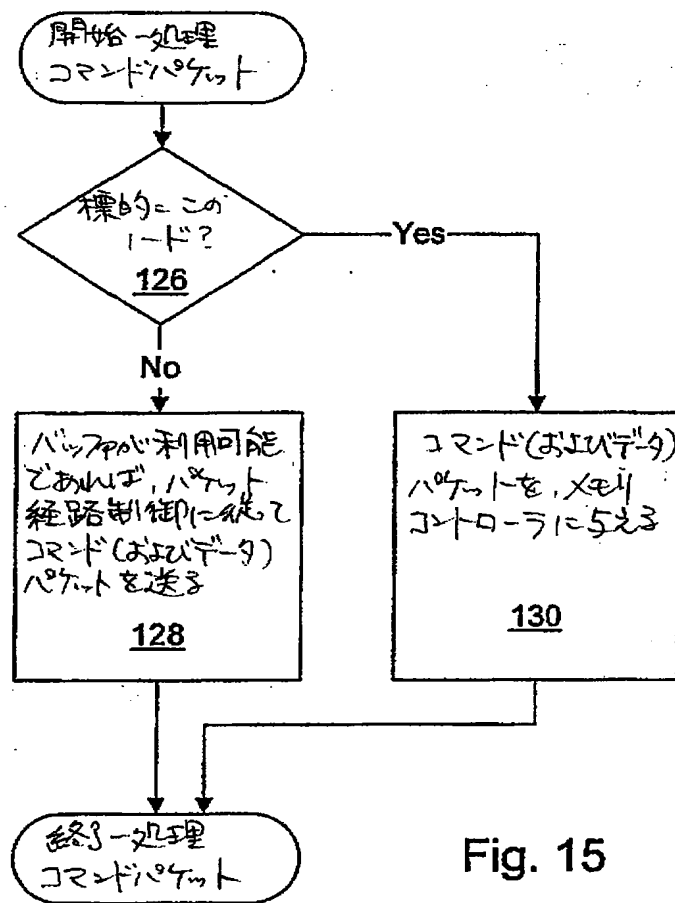


Fig. 15

【図16】

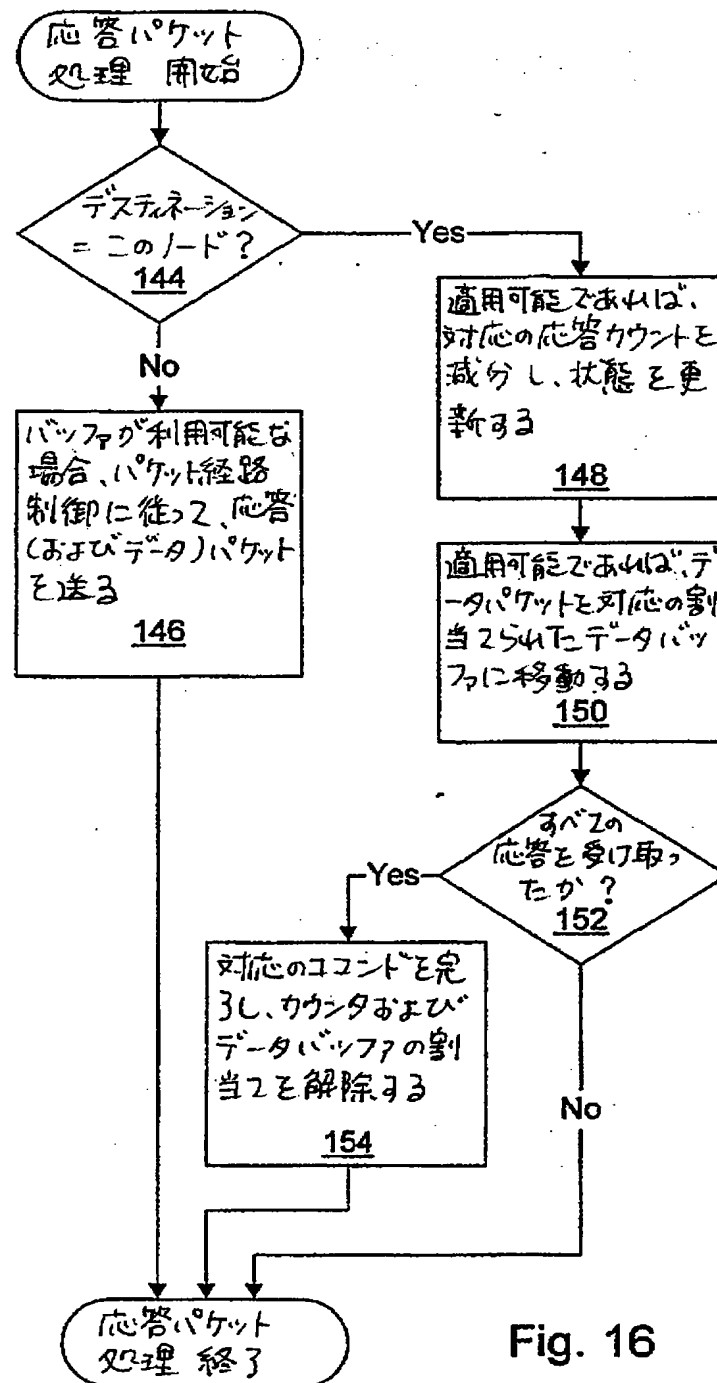


Fig. 16

【図17】

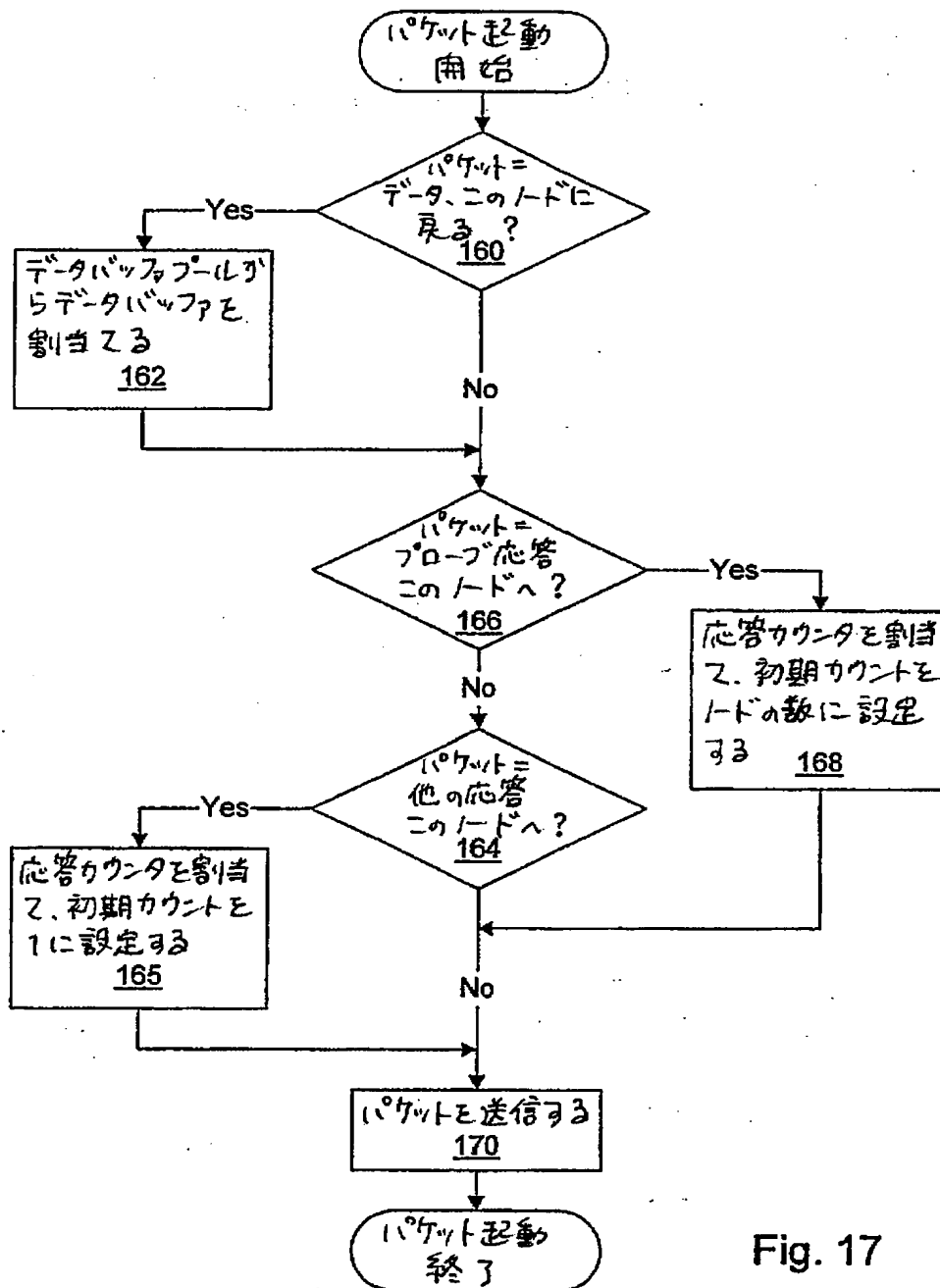


Fig. 17

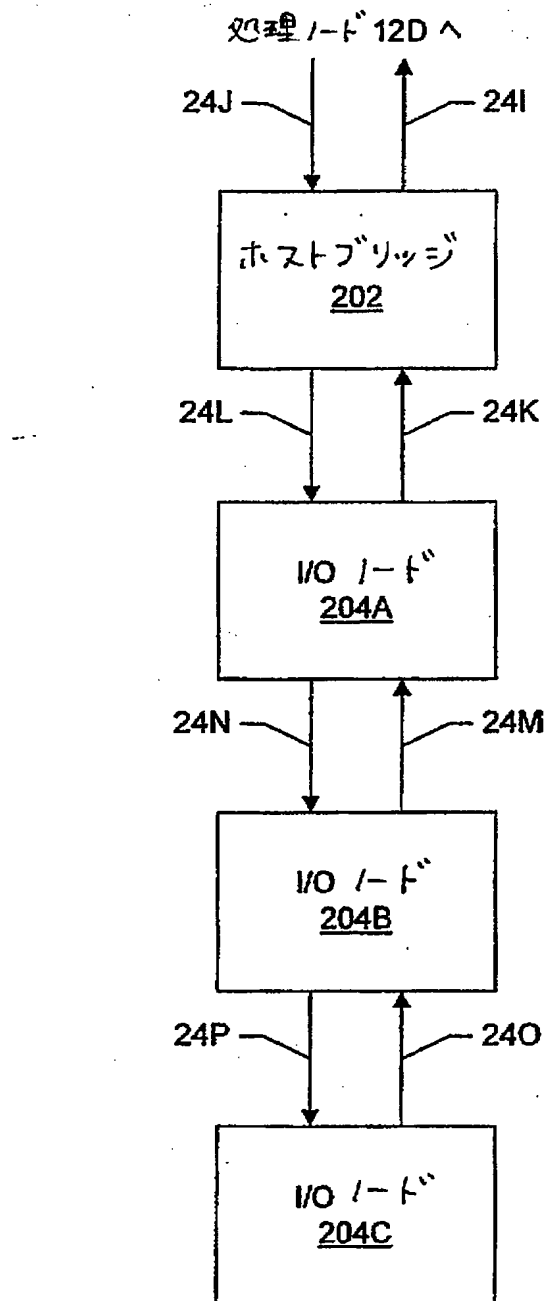
【図18】

ビット 時間	7	6	5	4	3	2	1	0
1	RSV		CMD[5:0]					
2	RespData [1:0]		Response [1:0]		RSV		RSV	
3	RSV		Probe [1:0]		Cmd Data[1:0]		Cmd[1:0]	
4	RSV							

180 

Fig. 18

【図19】



200

Fig. 19

【図20】

CMD コード ^a	VChan	コマンド ^b	パケットタイプ ^c
000000	-	Nop	Info
000001		Reserved	
000010	C	Flush	Command
000011		Reserved	
0001xx		Reserved	
x01xxx	C	WrSized	Command/Data
01xxxx	C	ReadSized	Command
100xxx		Reserved	
110000	R	RdResponse	Response/Data
110001		Reserved	
110010		Reserved	
110011	R	TgtDone	Response
11010x		Reserved	
110110		Reserved	
110111	R	AssignAck	Response
11100x		Reserved	
111010	C	Broadcast	Command
111011		Reserved	
11110x		Reserved	
111110	-	Assign	Info
111111	-	Sync	Info

210

Fig. 20

【図21】

ビット 時間	7	6	5	4	3	2	1	0
1	SeqID[3:2]		CMD[5:0]					
2		SeqID[1:0]		UnitID[4:0]				
3				SrcTag[4:0]				
4								
5	Addr[15:8]							
6	Addr[23:16]							
7	Addr[31:24]							
8	Addr[39:32]							


212 

Fig. 21

【図22】

ビット 時間	7	6	5	4	3	2	1	0
1			CMD[5:0]					
2				UnitID[4:0]				
3				SrcTag[4:0]				
4								


214 

Fig. 22

【図23】

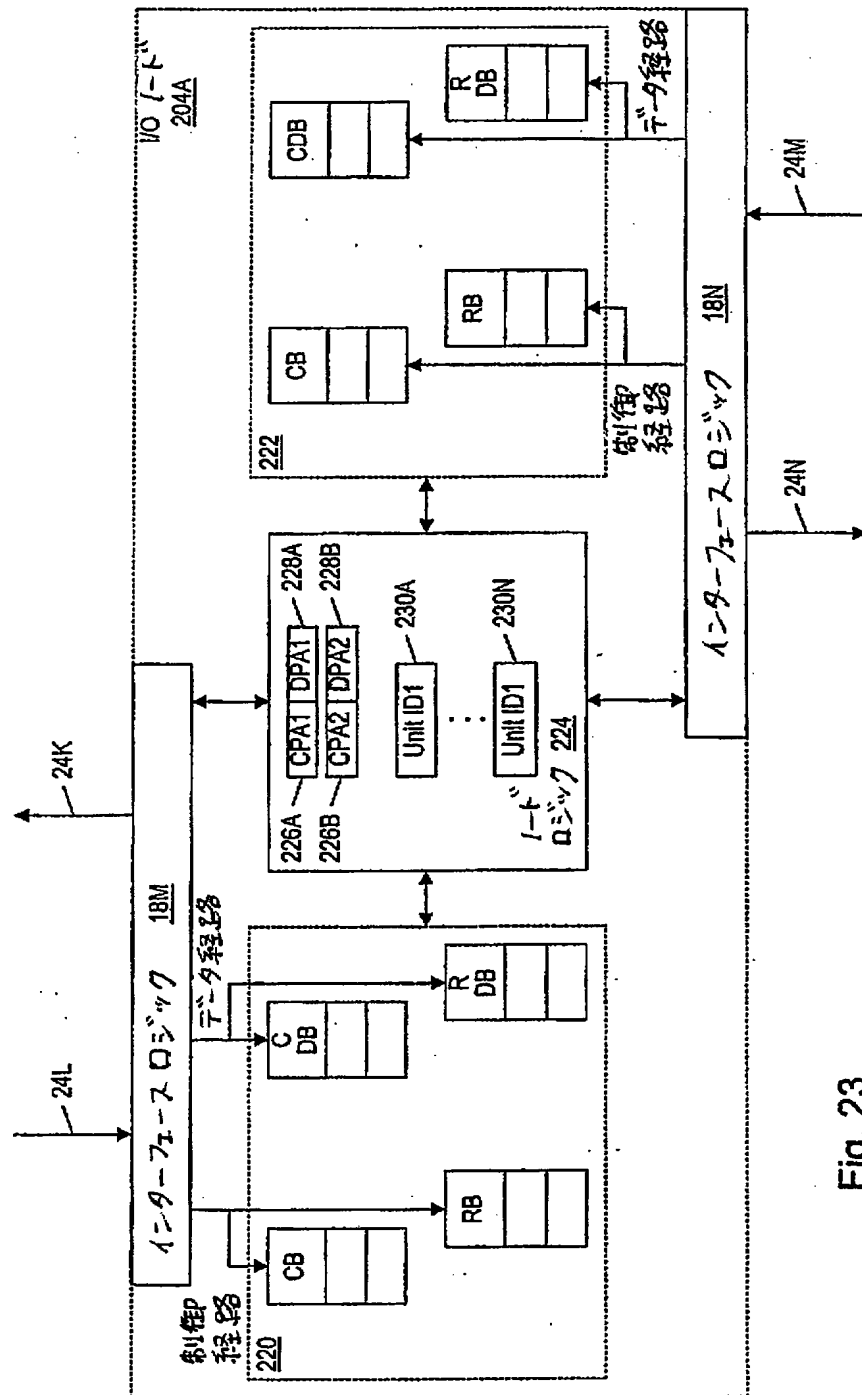


Fig. 23

【図24】

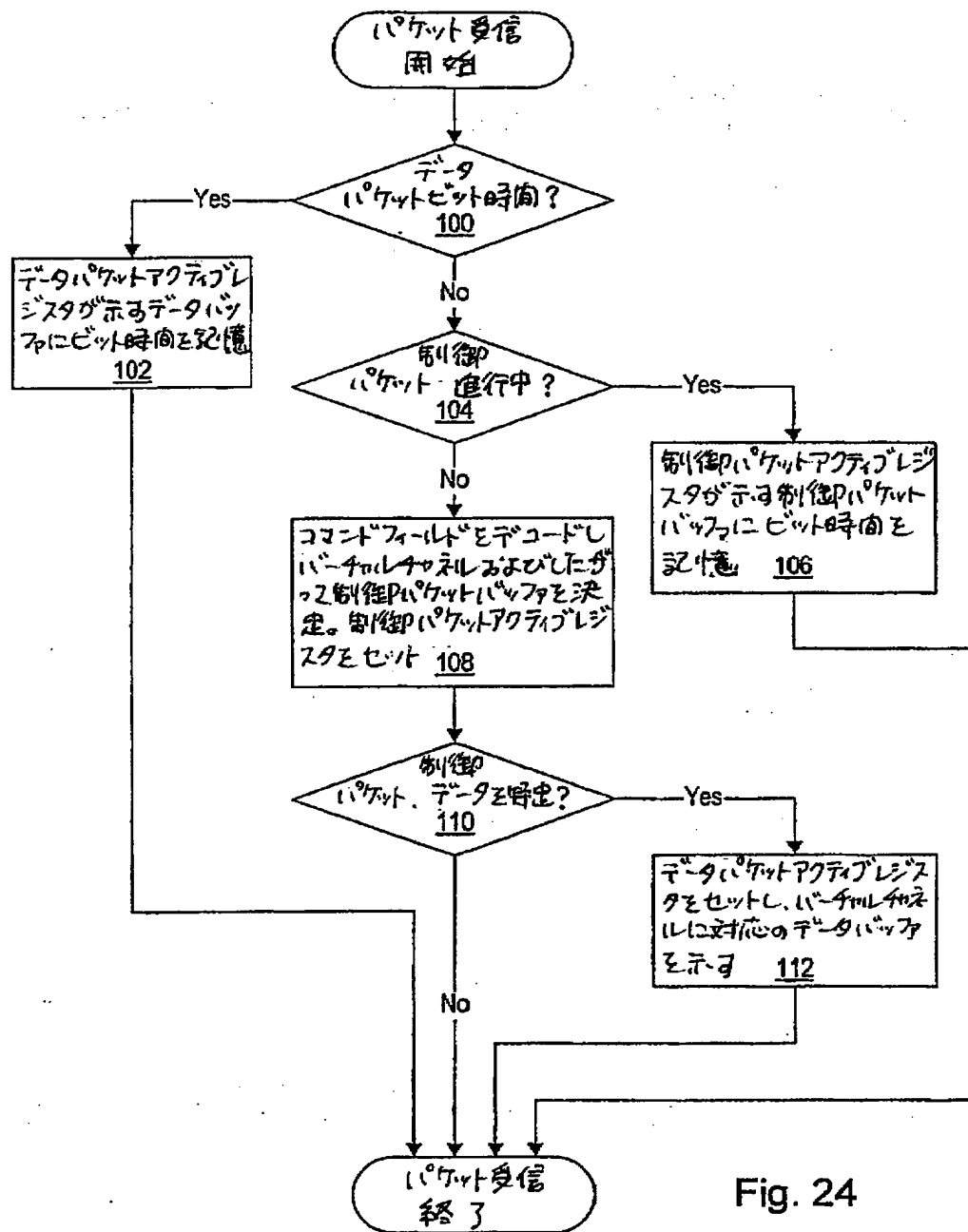


Fig. 24

【図25】

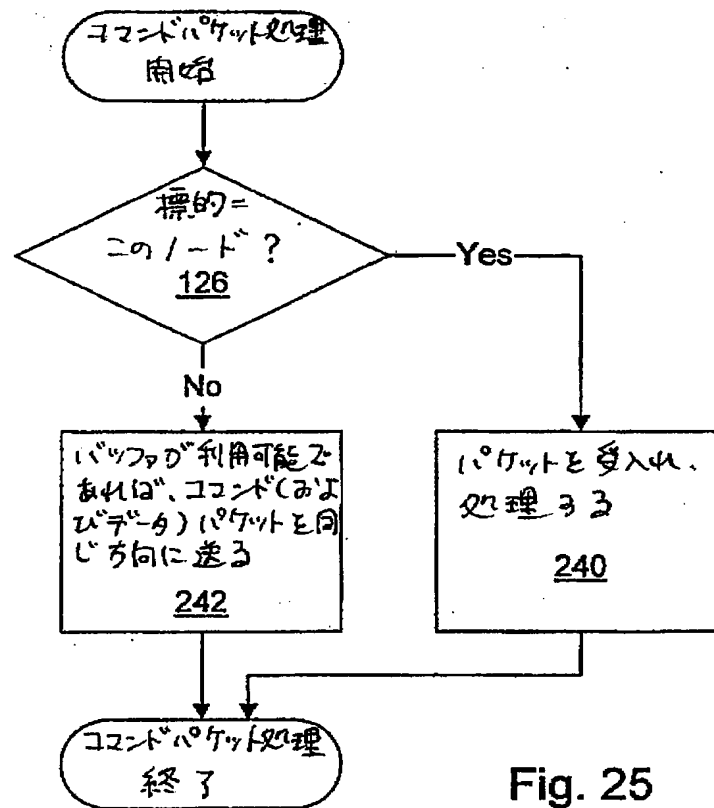


Fig. 25

【図26】

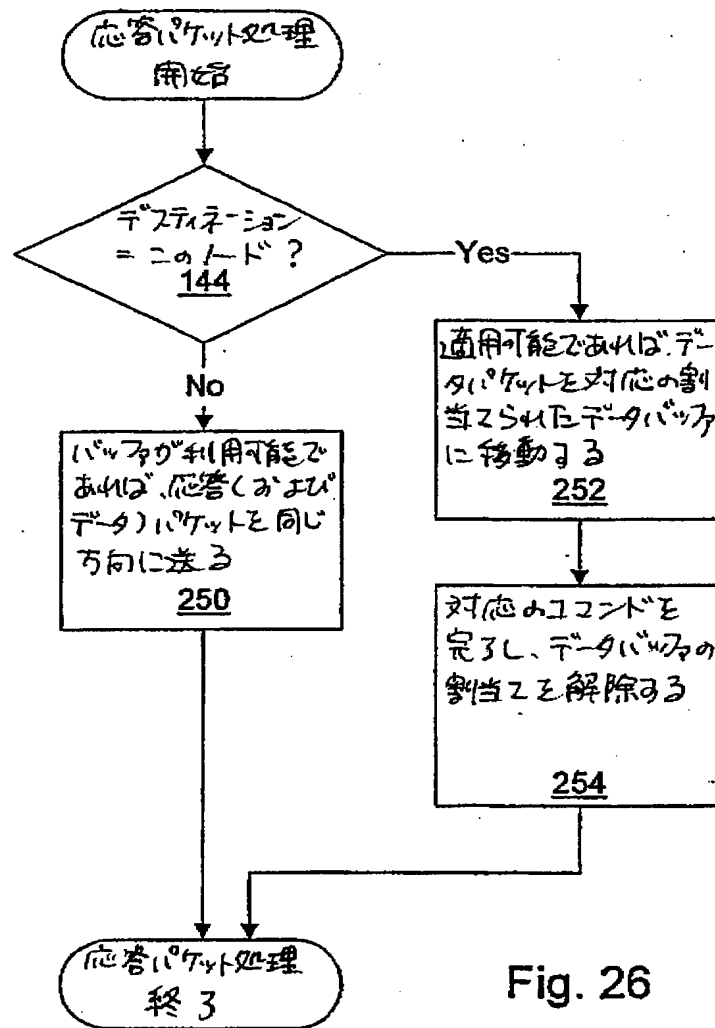


Fig. 26

【図27】

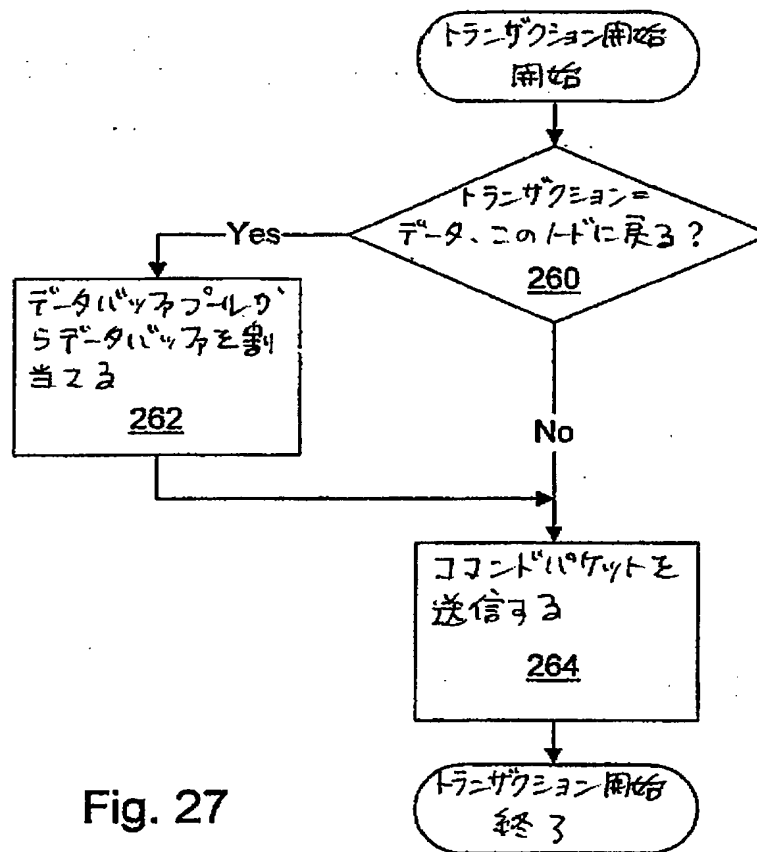


Fig. 27

【手続補正書】特許協力条約第34条補正の翻訳文提出書

【提出日】平成13年10月25日(2001.10.25)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】 パケットをコンピュータ(10)において複数のノード(12A～12D、204A～204C)間で経路づけるための方法であって、

第1の制御パケットを前記複数のノード(12A～12D、204A～204C)のうち第1のノード(12A、204A)において受取るステップを含み、前記第1のノード(12A、204A)は、複数のバーチャルチャネル(40A～40B)のうち異なるバーチャルチャネルに各々割当てられる複数のパケットバッファ(60A～60C、220)を含み、前記方法はさらに、

前記複数のバーチャルチャネル(40A～40B)のうち前記第1のパケットが属する第1のバーチャルチャネル(40A)を判断するステップと、

前記第1のパケットを前記第1のバーチャルチャネル(40A)に割当てられる第1のパケットバッファ(60A、220)において記憶するステップとを含み、前記方法は、

第1のパケットは制御パケットであることにおいて特徴付けられ、前記方法はさらに、

前記第1の制御パケットにより特定される第1のデータパケットを受取るステップと、

前記第1のデータパケットを前記第1のノード(12A、204A)内において複数のデータバッファ(60D～60E、220)のうち第1のデータバッファ(60D、220)において記憶し、前記複数のデータバッファ(60D～60E、220)の各々は、前記複数のバーチャルチャネル(40A～40B)のうち異なるバーチャルチャネルに割当てられる、方法。

【請求項2】 第1のパケットが属する第1のバーチャルチャネルの判断は第1のパケットのパケットタイプに依存する、請求項1に記載の方法。

【請求項3】 第2の制御パケットを前記第1のノード（12A、204A）において受取るステップと、

前記複数のバーチャルチャネル（40A～40B）のうち前記第2の制御パケットが属する第2のバーチャルチャネル（40B）を判断するステップとを含み、前記第2のバーチャルチャネル（40B）は前記第1のバーチャルチャネル（40A）とは異なり、前記方法はさらに、

前記第2の制御パケットを前記複数の制御パケットバッファ（60A～60C、220）のうち第2の制御パケットバッファ（60B、220）において記憶するステップを含み、前記第2の制御パケットバッファ（60B、220）は前記第2のバーチャルチャネル（40B）に割当てられる、請求項1に記載の方法。

【請求項4】 前記第1の制御パケットを前記複数のノード（12A～12D、204A～204C）のうち第2のノード（12B、204B）に送信し、前記第2のノード（12B、204B）は第2の複数の制御パケットバッファ（60A～60C、220）を含み、前記第2の複数の制御パケットバッファ（60A～60C、220）の各々は前記複数のバーチャルチャネル（40A～40B）のうち異なるバーチャルチャネルに割当てられ、前記送信は前記第2の複数の制御パケットバッファ（60A～60C、220）のうち前記第1の制御パケットを記憶する空間を含む第3の制御パケットバッファ（60A、220）に応答し、前記第3の制御パケットバッファ（60A、220）は前記第1のバーチャルチャネル（40A）に割当てられ、前記方法はさらに、

前記第2の複数の制御パケットバッファ（60A～60C、220）のうち前記第2の制御パケットを記憶する空間を含む第4の制御パケットバッファ（60B、220）に応答して前記第2の制御パケットを前記第2のノードに送信するステップを含み、前記第4の制御パケットバッファ（60B、220）は前記第2のバーチャルチャネル（40B）に割当てられる、請求項3に記載の方法。

【請求項5】 第1のパケットを送信するよう構成される第1のノード（1

2 A、2 0 4 B) と、

前記第1のパケットを前記第1のノード(1 2 A、2 0 4 B)から受信するよう結合される第2のノード(1 2 B、2 0 4 A)とを含み、前記第2のノード(1 2 B、2 0 4 A)は複数のパケットバッファ(6 0 A~6 0 C、2 2 0)を含み、前記複数のパケットバッファ(6 0 A~6 0 C、2 2 0)の各々は複数のバーチャルチャネル(4 0 A~4 0 B)のうち異なるバーチャルチャネルに割当てられ、前記第2のノード(1 2 B、2 0 4 A)は、前記複数のバーチャルチャネル(4 0 A~4 0 B)のうち前記第1のパケットが属する第1のバーチャルチャネル(4 0 A)に応答して前記複数のパケットバッファ(6 0 A~6 0 C、2 2 0)のうち第1のパケットバッファ(6 0 A、2 2 0)において前記第1のパケットを記憶するよう構成され、

第1のパケットは制御パケットであり、前記第2のノード(1 2 B、2 0 4 A)はさらに複数のデータバッファ(6 0 D~6 0 E、2 2 0)を含み、前記複数のデータバッファ(6 0 D~6 0 E、2 2 0)の各々は、前記複数のバーチャルチャネル(4 0 A~4 0 B)のうち、対応するデータパケットを特定する少なくとも1つの制御パケットを含む、異なるバーチャルチャネルに割当てられることを特徴とする、コンピュータシステム。

【請求項6】 第2のノードは第1のパケットのパケットタイプに依存して第1のパケットが属する第1のバーチャルチャネルを判断するよう構成される、請求項5に記載のコンピュータシステム。

【請求項7】 前記第1のノード(1 2 A、2 0 4 B)は前記第1の制御パケットにより特定される第1のデータパケットを送信するよう構成され、前記第2のノード(1 2 B、2 0 4 A)は前記第1のデータパケットを前記複数のデータバッファ(6 0 D~6 0 E、2 2 0)のうち第1のデータバッファ(6 0 D、2 2 0)に記憶するよう構成され、前記第1のデータバッファ(6 0 D、2 2 0)は前記第1のバーチャルチャネル(4 0 A)に割当てられる、請求項5または6に記載のコンピュータシステム。

【請求項8】 前記第1のノード(1 2 A、2 0 4 B)は前記複数のバーチャルチャネル(4 0 A~4 0 B)のうち第2のバーチャルチャネル(4 0 B)に

属する第2の制御パケットを送信するよう構成され、前記第2のバーチャルチャネル(40B)は前記第1のバーチャルチャネル(40A)とは異なり、前記第2のノード(12B、204A)は前記第2の制御パケットを前記複数の制御パケットバッファ(60A~60C、220)のうち第2の制御パケットバッファ(60B、220)において記憶するよう構成される、請求項5に記載のコンピュータシステム。

【請求項9】 前記第2のノード(12B、204A)が前記第1の制御パケットのデスティネーションである場合、前記第2のノード(12B、204A)は前記第1の制御パケットを前記第1の制御パケットバッファ(60A、220)から取除き前記第1の制御パケットに応答するよう構成され、前記第2のノード(12B、204A)は前記第1のデータパケットを前記第1のデータバッファ(60D、220)から取除き前記第1のデータパケットを処理するようさらに構成される、請求項7に記載のコンピュータシステム。

【請求項10】 パケットを前記第2のノード(12B、204A)から受信するよう結合される第3のノード(12C、204C)をさらに含み、前記第2のノード(12B、204A)が前記第1の制御パケットのデスティネーションでない場合、前記第2のノード(12B、204A)は前記第1の制御パケットを前記第1の制御パケットバッファ(60A、220)から取除き前記第1の制御パケットを前記第3のノード(12C、204C)へ送るよう構成され、前記第2のノード(12B、204A)は前記第1のデータパケットを前記第1のデータバッファ(60D、220)から取除き前記第1のデータパケットを前記第3のノード(12C、204C)に送るようさらに構成される、請求項7に記載のコンピュータシステム。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0007

【補正方法】変更

【補正の内容】

【0007】

V・S・アドヴェおよびM・K・ヴァーノン (V. S. Adve and M. K. Vernon) の「判断経路制御によるメッシュ相互接続ネットワークのパフォーマンス分析 (Performance Analysis of mesh interconnection networks with deterministic routing) INTERNET [Online] 1994年、1～40頁、XP2155424 は、物理的リンクの各々が、別々のバッファによって与えられる2つのバーチャルチャネルにより共用される、デッドロックのない経路制御アルゴリズムを用いた処理ノードのネットワークの間のメッセージを経路制御する方法を開示する。

上に概要を述べた問題は、ここで説明するようにバーチャルチャネルを用い、かつバーチャルチャネルに異なった資源を割当てることにより、大部分が解決される。

この発明の1つの局面によると、コンピュータシステム内の複数のノードの間でパケットを経路制御するための方法が開示され、方法は請求項1に定義される。

この発明の第2の局面によると、請求項5に定義されるコンピュータシステムが提供される。

一般的に、ロジック的／プロトコル関連のコンフリクトを有さないパケットは、バーチャルチャネルにグループ化される。したがって、ロジック的コンフリクトは別々のバーチャルチャネル内のパケットの間で生じる。1つのバーチャルチャネル内のパケットは資源を共用する（よって資源コンフリクトが生じる）が、異なったバーチャルチャネル内のパケットは資源を共用しないであろう。資源コンフリクトが生じるパケットはロジック的コンフリクトは生じなく、かつロジック的コンフリクトが生じるパケットは資源コンフリクトは生じないために、デッドロックのない動作が達成される。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0009

【補正方法】削除

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】 0 0 1 0

【補正方法】 削除

【国際調査報告】

INTERNATIONAL SEARCH REPORT

Internal Application No.
PCT/US 00/12574

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 606F15/173

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 606F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 533 198 A (G. M. THORSON) 2 July 1996 (1996-07-02) column 3, line 46 -column 9, line 29; claims; figures 5-7 ---	1-10
A	US 5 754 789 A (A. G. NOWATZYK) 19 May 1998 (1998-05-19) column 2, line 19 -column 13, line 34; claims; figures 9-12 --- -/-	1-10

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

A document member of the same patent family

Date of the actual completion of the international search

13 December 2000

Date of mailing of the international search report

05/01/2001

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2200 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 051 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Soler, J

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 00/12574

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>V. S. ADVE, M. K. VERNON: "Performance Analysis of mesh interconnection networks with deterministic routing" INTERNET, 'Online! 1994, pages 1-40, XP002155424</p> <p>Retrieved from the Internet: <URL:rice.edu/~adve/Papers...TPDS.mesh.ps.gz> 'retrieved on 2000-12-13! the whole document</p> <p>-----</p>	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 00/12574

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5533198 A	02-07-1996	EP 0671034 A JP 8503799 T WO 9412939 A US 5689646 A	13-09-1995 23-04-1996 09-06-1994 18-11-1997
US 5754789 A	19-05-1998	US 6081844 A	27-06-2000

フロントページの続き

- (72)発明者 ケラー, ジェイムス・ビィ
アメリカ合衆国、94303 カリフォルニア
州、パロ・アルト、アイリス・ウェイ、
210
- (72)発明者 メイヤー, デリック・アール
アメリカ合衆国、78733 テキサス州、オ
ースティン、クリークス・エッジ・パーク
ウェイ、2728
- F ターム(参考) 5B005 KK02 MM01 PP11
5B045 BB28 BB42 GG01 JJ40
5B077 AA13 DD02 NN01
5K033 AA02 BA04 CC02 DB12 DB14

【公報種別】特許法第 17 条の 2 の規定による補正の掲載
 【部門区分】第 6 部門第 3 区分
 【発行日】平成 19 年 5 月 31 日 (2007.5.31)

【公表番号】特表 2003-510686 (P2003-510686A)
 【公表日】平成 15 年 3 月 18 日 (2003.3.18)
 【出願番号】特願 2001-525542 (P2001-525542)
 【国際特許分類】

G O 6 F 15/177 (2006.01)
 G O 6 F 12/08 (2006.01)
 G O 6 F 13/38 (2006.01)
 H O 4 L 12/28 (2006.01)

【F I】

G O 6 F 15/177 6 7 6 A
 G O 6 F 15/177 6 7 8 G
 G O 6 F 12/08 5 3 1 B
 G O 6 F 13/38 3 1 0 C
 H O 4 L 12/28 1 0 0 C

【手続補正書】
 【提出日】平成 19 年 4 月 5 日 (2007.4.5)
 【手続補正 1】
 【補正対象書類名】明細書
 【補正対象項目名】特許請求の範囲
 【補正方法】変更
 【補正の内容】

【特許請求の範囲】

【請求項 1】 パケットをコンピュータシステムにおいて複数のノード間で経路づけるための方法であって、

第 1 の制御パケットを前記複数のノードのうち第 1 のノードにおいて受取ることを含み、前記第 1 のノードは複数の制御パケットバッファを含み、前記複数の制御パケットバッファの各々は複数のバーチャルチャネルのうち異なるバーチャルチャネルに割当てられ、前記方法はさらに、

前記複数のバーチャルチャネルのうち前記第 1 の制御パケットが属する第 1 のバーチャルチャネルを判断することと、

前記第 1 の制御パケットを前記複数の制御パケットバッファのうち第 1 の制御パケットバッファにおいて記憶することとを含み、前記第 1 の制御パケットバッファは前記第 1 のバーチャルチャネルに割当てられ、前記方法はさらに、

前記第 1 の制御パケットにより特定される第 1 のデータパケットを受取ることと、

前記第 1 のデータパケットを前記第 1 のノード内において複数のデータバッファのうち第 1 のデータバッファにおいて記憶することとを含み、前記複数のデータバッファの各々は、前記複数のバーチャルチャネルのうち、対応するデータパケットを特定する少なくとも 1 つの制御パケットを含む異なるバーチャルチャネルに割当てられる、方法。

【請求項 2】 第 2 の制御パケットを前記第 1 のノードにおいて受取ることと、

前記複数のバーチャルチャネルのうち前記第 2 の制御パケットが属する第 2 のバーチャルチャネルを判断することとを含み、前記第 2 のバーチャルチャネルは前記第 1 のバーチャルチャネルとは異なり、前記方法はさらに、

前記第 2 の制御パケットを前記複数の制御パケットバッファのうち第 2 の制御パケットバッファにおいて記憶することを含み、前記第 2 の制御パケットバッファは前記第 2 のバーチャルチャネルに割当てられる、請求項 1 に記載の方法。

【請求項 3】 前記第 1 の制御パケットを前記複数のノードのうち第 3 のノードに送信することをさらにを含み、前記第 3 のノードは第 2 の複数の制御パケットバッファを含み、前記第 2 の複数の制御パケットバッファの各々は前記複数のバーチャルチャネルのうち異なるバーチャルチャネルに割当てられ、前記送信は前記第 2 の複数の制御パケットバッファのうち前記第 1 の制御パケットを記憶する空間を含む第 3 の制御パケットバッファに回答し、前記第 3 の制御パケットバッファは前記第 1 のバーチャルチャネルに割当てられ、前記方法はさらに、

前記第 2 の複数の制御パケットバッファのうち前記第 2 の制御パケットを記憶する空間を含む第 4 の制御パケットバッファに回答して前記第 2 の制御パケットを前記第 3 のノードに送信することを含み、前記第 4 の制御パケットバッファは前記第 2 のバーチャルチャネルに割当てられる、請求項 2 に記載の方法。

【請求項 4】 前記判断することは、前記第 1 の制御パケットのコマンドフィールドをデコードすることを含む、請求項 1 に記載の方法。

【請求項 5】 前記判断することは、前記第 1 の制御パケットはポストされないコマンドバーチャルチャネルに属すると判断することを含む、請求項 1 に記載の方法。

【請求項 6】 前記判断することは、前記第 1 の制御パケットはプローブバーチャルチャネルに属すると判断することを含む、請求項 1 に記載の方法。

【請求項 7】 前記判断することは、前記第 1 の制御パケットは応答バーチャルチャネルに属すると判断することを含む、請求項 1 に記載の方法。

【請求項 8】 前記複数のバーチャルチャネルのうち少なくとも 1 つのバーチャルチャネルに含まれる各制御パケットはデータパケットを特定せず、前記複数のデータバッファのいずれも前記少なくとも 1 つのバーチャルチャネルに割当てられない、請求項 1 に記載の方法。

【請求項 9】 第 1 の制御パケットを送信するよう構成される第 1 のノードと、

前記第 1 の制御パケットを前記第 1 のノードから受信するよう結合される第 2 のノードとを含み、前記第 2 のノードは複数の制御パケットバッファを含み、前記複数の制御パケットバッファの各々は複数のバーチャルチャネルのうち異なるバーチャルチャネルに割当てられ、前記第 2 のノードは、前記複数のバーチャルチャネルのうち前記第 1 の制御パケットが属する第 1 のバーチャルチャネルに回答して前記複数の制御パケットバッファのうち第 1 の制御パケットバッファにおいて前記第 1 の制御パケットを記憶するよう構成され、前記第 2 のノードは複数のデータバッファをさらに含み、前記複数のデータバッファの各々は、前記複数のバーチャルチャネルのうち、対応するデータパケットを特定する少なくとも 1 つの制御パケットを含む、異なるバーチャルチャネルに割当てられ、前記第 1 のノードは前記第 1 の制御パケットにより特定される第 1 のデータパケットを送信するよう構成され、前記第 2 のノードは前記第 1 のデータパケットを前記複数のデータバッファのうち第 1 のデータバッファに記憶するよう構成され、前記第 1 のデータバッファは前記第 1 のバーチャルチャネルに割当てられる、コンピュータシステム。

【請求項 10】 前記第 1 のノードは前記複数のバーチャルチャネルのうち第 2 のバーチャルチャネルに属する第 2 の制御パケットを送信するよう構成され、前記第 2 のバーチャルチャネルは前記第 1 のバーチャルチャネルとは異なり、前記第 2 のノードは前記第 2 の制御パケットを前記複数の制御パケットバッファのうち第 2 の制御パケットバッファにおいて記憶するよう構成される、請求項 9 に記載のコンピュータシステム。

【請求項 11】 第 2 の複数の制御パケットバッファを含む第 3 のノードをさらに含み、前記第 2 の複数の制御パケットバッファの各々は前記複数のバーチャルチャネルのうち異なるバーチャルチャネルに割当てられ、前記第 2 のノードは前記第 2 の複数の制御パケットバッファのうち前記第 1 の制御パケットを記憶する空間を含む第 3 の制御パケットバッファに回答して前記第 1 の制御パケットを前記第 3 のノードに送信するよう構成され、前記第 3 の制御パケットバッファは前記第 1 のバーチャルチャネルに割当てられ、前記第 2 のノードは前記第 2 の複数の制御パケットバッファのうち前記第 2 の制御パケットを記憶する空間を含む第 4 の制御パケットバッファに回答して前記第 2 の制御パケットを前

記第 3 のノードに送信するよう構成され、前記第 4 の制御パケットバッファは前記第 1 のバーチャルチャネルに割当てられる、請求項 10 に記載のコンピュータシステム。

【請求項 12】 前記第 2 のノードが前記第 1 の制御パケットのデスティネーションである場合、前記第 2 のノードは前記第 1 の制御パケットを前記第 1 の制御パケットバッファから取除き前記第 1 の制御パケットに回答するよう構成される、請求項 9 に記載のコンピュータシステム。

【請求項 13】 前記第 2 のノードは、さらに、前記第 1 のデータパケットを前記第 1 のデータバッファから取除き前記第 1 のデータパケットを処理するよう構成される、請求項 12 に記載のコンピュータシステム。

【請求項 14】 前記第 2 のノードはキャッシュおよびメモリコントローラを含み、前記第 2 のノードは前記第 1 の制御パケットに回答して前記第 1 のデータパケットを前記キャッシュおよび前記メモリコントローラのうちの 1 つに与えるよう構成される、請求項 13 に記載のコンピュータシステム。

【請求項 15】 パケットを前記第 2 のノードから受信するよう結合される第 3 のノードをさらに含み、前記第 2 のノードが前記第 1 の制御パケットのデスティネーションでない場合、前記第 2 のノードは前記第 1 の制御パケットを前記第 1 の制御パケットバッファから取除き前記第 1 の制御パケットを前記第 3 のノードへ送るよう構成される、請求項 9 に記載のコンピュータシステム。

【請求項 16】 前記第 2 のノードは前記第 1 のデータパケットを前記第 1 のデータバッファから取除き前記第 1 のデータパケットを前記第 3 のノードに送るようさらに構成される、請求項 14 に記載のコンピュータシステム。

【請求項 17】 前記第 2 のノードは、前記第 1 の制御パケットが属する前記第 1 のバーチャルチャネルを、前記第 1 の制御パケットのコマンドフィールドをデコードすることにより判断するよう構成される、請求項 9 に記載のコンピュータシステム。

【請求項 18】 前記複数のバーチャルチャネルのうち少なくとも 1 つのバーチャルチャネルに含まれる各制御パケットはデータパケットを特定せず、前記複数のデータバッファのいずれも前記少なくとも 1 つのバーチャルチャネルに割当てられない、請求項 9 に記載のコンピュータシステム。

【請求項 19】 第 1 の制御パケットおよび前記第 1 の制御パケットにより特定される第 1 のデータパケットを受取るよう結合されるノードであって、

複数の制御パケットバッファを含み、前記複数の制御パケットバッファの各々は複数のバーチャルチャネルのうちの異なるバーチャルチャネルに割当てられ、前記ノードはさらに、

複数のデータバッファを含み、前記複数のデータバッファの各々は、前記複数のバーチャルチャネルのうち、対応するデータパケットを特定する少なくとも 1 つの制御パケットを含む、異なるバーチャルチャネルに割当てられ、前記ノードはさらに、

前記複数のバーチャルチャネルのうち前記第 1 の制御パケットが属する第 1 のバーチャルチャネルに回答して前記第 1 の制御パケットを前記複数の制御パケットバッファのうちの第 1 の制御パケットバッファに記憶するよう構成され、さらに、前記第 1 のデータパケットを前記複数のデータバッファのうちの第 1 のデータバッファに記憶するよう構成される回路系を含み、前記第 1 のデータバッファは前記第 1 のバーチャルチャネルに割当てられる、ノード。

【請求項 20】 前記ノードが前記第 1 の制御パケットのデスティネーションである場合、前記回路系は前記第 1 の制御パケットを前記第 1 の制御パケットバッファから取除くよう構成され、前記ノードは前記第 1 の制御パケットに回答するよう構成される、請求項 19 に記載のノード。

【請求項 21】 前記回路系は、さらに、前記第 1 のデータパケットを前記第 1 のデータバッファから取除くよう構成され、前記ノードは前記第 1 のデータパケットを処理するよう構成される、請求項 20 に記載のノード。

【請求項 22】 キャッシュおよびメモリコントローラをさらに含み、前記ノードは

前記第 1 の制御パケットに応答して前記第 1 のデータパケットを前記キャッシュおよび前記メモリコントローラのうちの 1 つに与えるよう構成される、請求項 21 に記載のノード。

【請求項 23】 前記第 2 のノードが前記第 1 の制御パケットのデスティネーションでない場合、前記回路系は前記第 1 の制御パケットを前記第 1 の制御パケットバッファから取除き前記第 1 の制御パケットを別のノードへ送るよう構成される、請求項 19 に記載のノード。

【請求項 24】 前記回路系は前記第 1 のデータパケットを前記第 1 のデータバッファから取除き前記第 1 のデータパケットを前記別のノードに送るようさらに構成される、請求項 23 に記載のノード。

【請求項 25】 前記回路系は、前記第 1 の制御パケットが属する前記第 1 のバーチャルチャネルを、前記第 1 の制御パケットのコマンドフィールドをデコードすることにより判断するよう構成される、請求項 19 に記載のノード。

【請求項 26】 前記複数のバーチャルチャネルのうち少なくとも 1 つのバーチャルチャネルに含まれる各制御パケットはデータパケットを特定せず、前記複数のデータバッファのいずれも前記少なくとも 1 つのバーチャルチャネルに割当てられない、請求項 19 に記載のノード。